THE SEQUENTIAL IMPLEMENTATION OF ARRAY PROCESSORS WHEN THERE IS DIRECTIONAL UNCERTAINTY

William S. Hodgkiss, Jr.

Duke University

Prepared for:

Office of Naval Research

August 1975

Unclassified

Security Classification

DOCUMENT CONTROL DATA - R & D

(Security classification of title, body of abstract and indexing annotation must be entered when the overall report is classified)

1. ORIGINATING ACTIVITY (Corporate author)	2a. REPORT SECURITY CLASSIFICATION
Department of Electrical Engineering Duke University Durham, NC 27706	Unclassified 2b. GROUP

3. REPORT TITLE

The Sequential Implementation of Array Processors when there is Directional Uncertainty

4. DESCRIPTIVE NOTES (Type of report and inclusive dates)
Summary 1974-1975

5. AUTHOR(S) (First name, middle initial, last name)

William S. Hodgkiss, Jr.

6. REPORT DATE	7a. TOTAL NO. OF PAGES	7b. NO. OF REFS
August 1975	~~214~~ 235	34

8a. CONTRACT OR GRANT NO.	9a. ORIGINATOR'S REPORT NUMBER(S)
N00014-75-C-0191 b. PROJECT NO.	Adaptive Signal Detection Laboratory
c.	9b. OTHER REPORT NO(S) (Any other numbers that may be assigned this report) TR-10
d.	

10. DISTRIBUTION STATEMENT

Distribution of this document is unlimited.

11. SUPPLEMENTARY NOTES	12. SPONSORING MILITARY ACTIVITY
	Office of Naval Research Navy Department (Code 222)

13. ABSTRACT

This report takes a global approach to the processing of information from an array of sensors. Specifically, the detection problem is considered. The general form of the likelihood ratio is derived based upon observables consisting of the Fourier coefficients of the observed random processes. For a stationary noise field consisting of a component independent from sensor to sensor and an additive directional component, the covariance properties of these Fourier coefficients are pursued as a function of the observation period length.

Four canonical implementations of the optimal array processor structure are discussed: (1) one shot, (2) pseudo estimator, (3) two step, and (4) sequential. The pseudo estimator structure is shown to be the optimal counterpart of an appealing ad hoc approach to array processor design where any uncertain parameters are first estimated, then plugged into the parameters known likelihood ratio as if they were known exactly. The general formulation of the time sequential structure reveals that the likelihood ratio can be realized by an appropriate combination of single frequency components. It is shown that the optimal array processor exhibits learning or adaptive features naturally when implemented sequentially.

Of particular interest are three specific problems involving either signal or noise source location uncertainty. Their likelihood ratio expressions are derived and performance reported for several levels of location uncertainty and two array sizes. Performance is stated in terms of the ROC curve.

i.

DD FORM 1 NOV 68 1473

Unclassified

Security Classification

PAGES _____

ARE

MISSING

IN

ORIGINAL

DOCUMENT

14 KEY WORDS	LINK A		LINK B		LINK C	
	ROLE	WT	ROLE	WT	ROLE	WT
Optimum Signal Detection Theory						
Bayesian						
Likelihood Ratio						
Detection Performance						
Receiver Structure						
Sequential Processing						
Adaptive Array Processors						
Signal Representation						
Sufficient Statistics						
Uncertain Location						
Gaussian Signal						
Signal Known Exactly						
Directional Noise Source						
Underwater Acoustics						
Acoustic Signal Processing						
Adaptive Arternas						

ie.

School of

ENGINEERING

DUKE UNIVERSITY

TR-10

THE SEQUENTIAL IMPLEMENTATION OF ARRAY PROCESSORS WHEN THERE IS DIRECTIONAL UNCERTAINTY

by
William S. Hodgkiss, Jr.
Department of Electrical Engineering

Prepared under:
Office of Naval Research (Code 222)
Contract No. N000-14-75-C-0191

DUKE UNIVERSITY

ADAPTIVE SIGNAL DETECTION LABORATORY

Department of Electrical Engineering
School of Engineering

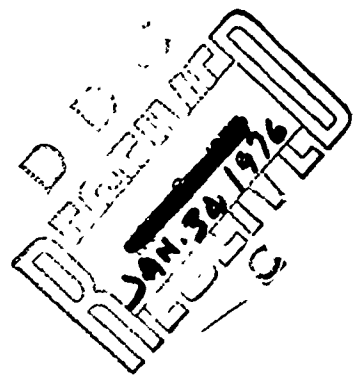
Technical Report No. 10

The Sequential Implementation of Array Processors when there is Directional Uncertainty

by

William S. Hodgkiss, Jr.

August 1975



Approved: Loren W Nolte
Loren W. Nolte
Principal Investigator

Prepared under: Office of Naval Research (Code 222)
Contract No. N00014-75-C-0191

This report was also a dissertation submitted in partial fulfillment of the requirements for the Doctor of Philosophy degree in the Department of Electrical Engineering in the Graduate School of Arts and Sciences of Duke University, 1975.

ABSTRACT

This dissertation takes a global approach to the processing of information from an array of sensors. Essentially, the unprocessed outputs of the individual elements are considered as the observables. The processor structure is allowed to evolve freely with the sole restriction being the criterion of optimality.

Specifically, the array processors discussed must decide if the random processes observed at the array element outputs consist of a signal obscured by noise or noise alone. Any uncertain parameters are treated as random variables and knowledge about them is summarized by a priori probability density functions. The resulting detectors are optimum in the sense of making a least-risk decision.

The general form of the likelihood ratio is derived based upon observables consisting of the Fourier coefficients of the observed random processes. For a stationary noise field consisting of a component independent from sensor to sensor and an additive directional component, the covariance properties of these Fourier coefficients are pursued as a function of the observation period length.

Once the mathematics of the likelihood ratio has been written, the optimal array processor can be implemented in various structures. Four such canonical implementations are discussed: (1) one shot, (2) pseudo estimator, (3) two step, and (4) sequential. The pseudo estimator structure is shown to be the optimal counterpart of an appealing ad hoc approach to array processor design where any uncertain parameters are

first estimated, then plugged into the parameters known likelihood ratio as if they were known exactly. The general formulation of the time sequential structure reveals that the likelihood ratio can be realized by an appropriate combination of single frequency components. Each is an independent time sequential processor which utilizes its own natural conjugate prior to achieve a certain degree of mathematical tractability.

Of particular interest are three specific problems involving either signal or noise source location uncertainty

(1) Gaussian Signal of Uncertain Direction in Gaussian Noise (GUD)

(2) Signal Known Except for Direction in Noise with an Additive Directional Noise Component of Known Direction (SKED in NKD)

(3) Signal Known Exactly in Noise with an Additive Directional Noise Component of Uncertain Direction (SKE in NUD).

Their likelihood ratio expressions are derived and performance reported for several levels of location uncertainty and two array sizes. Performance is stated in terms of the ROC curve.

Although an estimate and plug structure is appealing due to its explicit adaptive characteristics, it is shown that the optimal array processor exhibits learning or adaptive features naturally when implemented sequentially. Computer simulation runs of the SKE in NUD processor are used to illustrate the Bayesian updating which occurs as an integral part of the sequential structure.

ACKNOWLEDGEMENTS

The author wishes to express his sincere appreciation for the guidance and encouragement given by his advisor and doctoral committee chairman, Dr. L. W. Nolte. The support of the Office of Naval Research is also gratefully acknowledged. During a major portion of this research, the Department of Electrical Engineering at the University of Washington kindly supplied office space and computing facilities.

The author has benefited greatly from discussions with several other individuals. Particularly valuable has been the interaction with the remaining members of the Adaptive Signal Detection Laboratory and a former member, Dr. S. L. Adams.

Thanks are due to Ms. Barbara Hill of the University of Washington for the typing of this manuscript.

CONTENTS

Contents (continued)

Contents (continued)

LIST OF TABLES

LIST OF FIGURES

List of Figures (continued)

List of Figures (continued)

GLOSSARY OF SYMBOLS

A_0	parameter in the a priori probability density function summarizing location uncertainty
AO	A_0
A_ℓ	magnitude of the ℓ^{th} complex observation statistic
B_0	parameter in the a priori probability density function summarizing location uncertainty
BO	B_0
$b_0(n)$	Fourier coefficient at $\omega = n\omega_0$ of the deterministic signal observed at the 0^{th} array element output
B_ℓ	phase of the ℓ^{th} complex observation statistic
C	a constant
C_ℓ	complex statistic of the observation vector
D	height of the directional noise power spectrum
$D(n\omega_0)$	power spectral density function of the directional noise
$d_k(t)$	directional noise component at the k^{th} array element output
d^2	detectability index
D_0, D_1	processor responses for the binary decision problem
D_ℓ	complex statistic of the observation vector
E	received signal energy
$E[\cdot]$	expected value
$E[\cdot\|\cdot]$	conditional expected value
E_ℓ	complex statistic of the observation vector

Glossary of Symbols (continued)

$\text{erfc}_*(\cdot)$	error function complement
$\exp(\cdot)$	exponential function
$\exp(jx)$	$\cos(x) + j \sin(x)$ (Euler's formula)
F_ℓ	complex statistic of the observation vector
G_ℓ	complex statistic of the observation vector
$g[\underline{\delta}(\cdot),\underline{\theta}]$	function which depends on the observation vector only through $\underline{\delta}(\cdot)$
$G(\cdot)$	function which does not depend on $\underline{\theta}$
H_0, H_1	the two hypotheses in the binary decision problem
H0	H_0
H1	H_1
H_ℓ	complex statistic of the observation vector
$h(\underline{\theta};\underline{\gamma})$	member of a family of probability density functions indexed by the parameter vector $\underline{\gamma}$
$\mathcal{H}_\Gamma(\Theta)$	a family of probability density functions on Θ
I_ℓ	complex statistic of the observation vector
$I_0(\cdot)$	modified Bessel function of order zero
$\text{Im}\{\cdot\}$	imaginary part
ITER	iteration
j	imaginary unit
K	number of array elements
k	array element index

Glossary of Symbols (continued)

L	$\Lambda(\cdot)$ (likelihood ratio)
$\ell(\cdot)$	$\ln\Lambda(\cdot)$
ln $(\cdot)$	natural logarithm
lim	limit
l.i.m.	limit in the mean
$\underline{m}$	$E[\underline{z}]$ (mean or expected value of the vector $\underline{z}$)
$m(t_i)$	i^{th} time sample of the function modulating $\cos(\omega t-\phi)$
N	total number of Fourier coefficients less one; also used as the height of the independent noise power spectrum
$N(n\omega_c)$	power spectral density function of the noise which is independent from sensor to sensor
$N_0/2$	power spectrum height for white Gaussian noise
n	frequency index
$\underline{n}(\cdot)$	vector of all Fourier coefficients due to noise alone
$n_k(t)$	independent noise component at the k^{th} array element output
$N(0,\sigma^2)$	normal or Gaussian distribution with mean zero and variance σ^2
$p(\cdot)$	joint probability density function
$p(\cdot\|\cdot)$	conditional joint probability density function
$p'(\underline{\theta};\underline{\gamma})$	natural reproducing probability density function
P(WOTN)	$p(\omega_0\tau_n)$
P_D	detection probability
P_F	false alarm probability
PI	π
$\underline{Q}$	covariance matrix
$\underline{Q}^{-1}$	inverse of $\underline{Q}$
$\mathcal{R}$	general observation space
$\underline{R}$	general vector of observation
$\underline{R}_K$	general observation vector of dimension K

Glossary of Symbols (continued)

R_i	i^{th} observation
R^m	real vector space of dimension m
$R(\tau)$	autocorrelation function
$r(\underline{\theta})$	a nonnegative function defined on Θ
$Re\{\cdot\}$	real part
S	height of the random signal power spectrum
$S(n\omega_0)$	power spectral density function of the random signal
$\underline{s}(\cdot)$	vector of all Fourier coefficients due to signal alone
$s(n)$	the Fourier coefficient at $\omega = n\omega_0$ of a deterministic time waveform
$s(t)$	a deterministic time waveform
$s(t_i)$	i^{th} time sample of $s(t)$
$\text{sinc}(x)$	$\sin(\pi x)/(\pi x)$
T	total or incremental observation period length depending on the context; also used to indicate transpose
t	the time variable
T_{inc}	incremental observation period length (Chapter V)
U_ℓ	parameter in the conjugate a priori probability density function for the uncertain direction problems in Chapter VI
$\underline{u}(n,\tau_s)$	pointing vector in the direction of the signal source at $\omega = n\omega_0$
$(1/2)\underline{V}$	variance of both $\underline{x}$ and $\underline{y}$
V_ℓ	parameter in the conjugate a priori probability density function for the uncertain direction problems in Chapter VI
$\text{var}(\cdot)$	variance
$\underline{v}(n,\tau_n)$	pointing vector in the direction of the noise source at $\omega = n\omega_0$

Glossary of Symbols (continued)

$\underline{z}(t)$	random processes observed at all the array element outputs
$z_k(t)$	random process observed at the k^{th} array element output
Γ	parameter space
$\underline{\gamma}$	parameter vector characterizing one member of a family of probability density functions
$\underline{\gamma}^0$	<u>a priori</u> parameter vector
$\underline{\gamma}^K$	<u>a posteriori</u> parameter vector after K observations
$\underline{\delta}(\cdot)$	sufficient statistic
δ_{nm}	$\begin{cases} 1, & n = m \\ 0, & n \neq m \end{cases}$ (Kronecker delta function)
η	decision threshold
Θ	parameter space
$\underline{\theta}$	vector of parameters
$\Lambda(\cdot)$	likelihood ratio
$\Lambda(\cdot\|\cdot)$	conditional likelihood ratio
$\Lambda'(\cdot)$	likelihood ratio based on natural conjugate priors
σ^2	variance
τ	time delay
τ_s	time delay of the signal between adjacent elements of a uniformly spaced linear array
τ_n	time delay of the directional noise between adjacent elements of a uniformly spaced linear array
$\tau_{\ell k}$	time delay of the directional noise between the ℓ^{th} and k^{th} array elements
Φ	one-dimensional parameter space in the SKEP problem
ϕ	uncertain parameter (phase) in the SKEP problem

Glossary of Symbols (continued)

ω	the radian frequency variable
ω_0	$2\pi/T$ or $2\pi/T_{inc}$ (fundamental radian frequency)
$\lvert\cdot\rvert^2$	magnitude squared of the complex argument
$\sum_{n=0}^{N}$	summation
$\prod_{n=0}^{N}$	multiplication
$\int_{\Theta}(\cdot)d\underline{\theta}$	multiple integration over the parameter space
$\approx$	approximately equal to
$\rightarrow$	approaches
$\Longleftrightarrow$	Fourier transform pair
$\in$	membership
$\subset$	inclusion
$*$	conjugate transpose
$\hat{}$	circumflexes denote estimates of the true values

THE SEQUENTIAL IMPLEMENTATION OF ARRAY PROCESSORS WHEN THERE IS DIRECTIONAL UNCERTAINTY

Chapter I

INTRODUCTION AND OVERVIEW

The utilization of an array of sensors to obtain some degree of spatial filtering or directional sensitivity is well known. Most of the antenna array literature concerns itself with beam patterns achieved through various weightings and location geometries of the individual elements. Thus, it is not surprising that the approach to processing information from an array of sensors usually has involved beamforming as a basis. While such an approach seems quite logical, it inherently assumes that operations which appear correct locally (beamforming, for instance) will facilitate the overall goal of good signal processing.

A global approach to the processing of information from an array of sensors takes another point of view. Essentially, the unprocessed outputs of the individual elements are considered as the observables. Based upon some criterion of optimality, the processor structure is allowed to evolve freely out of the mathematics of the problem being considered. Beamforming may be an integral part of the resulting structure, but it is not imposed from the beginning.

The array processors discussed in this dissertation are derived from within such a global framework. Specifically, they must decide if the random processes observed at the array element outputs consist of a signal obscured by noise or noise alone. Any uncertain parameters in the problems considered are treated as random variables and knowledge about them is

summarized by a priori probability density functions. The resulting detectors are optimum in the sense of making a least-risk decision. This approach has become known as the Bayesian design philosophy.

Major contributions of this work are in two areas. First, the time sequential optimal array processor is formulated in general terms and its natural adaptive feature is noted. Second, the optimal array processors are derived and their performance reported for three specific problems involving either signal or noise source location uncertainty.

Detection theory from a Bayesian point of view is reviewed in Chapter II. The likelihood ratio is introduced as the optimum processor for any goodness criterion where good decisions are preferred over bad. Then, four canonical structures in which the likelihood ratio may be implemented are discussed. Lastly, evaluation of performance for optimal processors is considered.

The general signal detection theory results of Chapter II are specialized in Chapter III to the optimal processing of data from an array of sensors. A review of the related literature is presented followed by a discussion on the utilization of the Fourier coefficients of the observed random processes as observables for an optimal array processor. The likelihood ratio based on these Fourier coefficients is derived and its implementation in terms of three of the canonical structures mentioned above is discussed. An important comparison is made between the optimal and a popular ad hoc approach to the estimate and plug array processor.

In preparation for a discussion on the time sequential implementation of the optimal array processor, Chapter IV considers the covariance between the Fourier coefficients as a function of the observation period length.

Specific results for a noise field consisting of a component independent from sensor to sensor and an additive directional component are given.

Chapter V presents the general formulation of the time sequential optimal array processor. A canonical structure is derived where it is shown that the likelihood ratio may be realized by an appropriate combination of single frequency components. For the particular noise field described above, conditions are given for the selection of an incremental observation period length.

Three specific detection problems where there is either signal or noise source location uncertainty are the subject of Chapter VI. Their parameter conditional joint density expressions are derived and the essential features of each is illustrated.

A generalized approach to the analysis of performance for optimal processors is discussed in Chapter VII. The utilization of sufficient statistics as an intermediate step between the observables and the likelihood ratio is shown to be advantageous when performance for more than one pair of a prior probability density functions is desired.

Chapter VIII contains a presentation and detailed discussion of performance results for the three specific problems mentioned above. Several observations comparing the relative seriousness of signal and noise source location uncertainty are made.

The natural adaptive feature of an optimal array processor when implemented sequentially is considered again in Chapter IX. Computer simulation runs where the directional noise source's location is uncertain are used to illustrate the Bayesian updating which occurs as an integral part of the sequential structure. Several noise-to-noise ratios are investigated.

Lastly, Chapter X provides a summary of this work and suggestions for further research.

Chapter II

BAYESIAN SIGNAL DETECTION THEORY

Within the communications context, the two broad subjects of signal detection and estimation theory are concerned with decision making based upon operations performed on some received data. In the first, only a decision about the presence or absence of a certain signal, or subset of signals, in the data is required. In the second, the decision involves estimating one or several parameters which are contained in the data. Processors for these tasks are designed based on some criterion of goodness or optimality. The viewpoint in this dissertation will be Bayesian where any uncertain parameters are modeled as random variables and knowledge about them is summarized by a priori probability density functions.

The following sections will provide the mathematical framework of signal detection theory from which we will work. Important concepts to be emphasized are: (1) the likelihood ratio, (2) processor structure, and (3) performance. Parts of the presentation will closely follow a recent excellent paper by Birdsall and Gobien (Birdsall and Gobien, 1973).

The Likelihood Ratio

Consider the binary decision problem where there are two mutually exclusive and exhaustive hypotheses, H_0 and H_1. Assume a vector of observations $\underline{R}$ is made from a space $\mathcal{R}$. Under H_0, the distribution on $\mathcal{R}$ is characterized by a probability density conditioned on a vector of parameters $\underline{\theta}_0$ which belongs to the family $\{p(\underline{R}|\underline{\theta}_0,H_0)\ ;\ \underline{\theta}_0 \in \Theta_0\}$. Under H_1,

the density belongs to the family $\{p(\underline{R}|\underline{\theta}_1, H_1) \; ; \; \underline{\theta}_1 \in \Theta_1\}$. There may be components of $\underline{\theta}_0$ and $\underline{\theta}_1$ which represent the same parameters. In summary

$$\begin{aligned} H_0 &: p(\underline{R}|\underline{\theta}_0, H_0) \quad , \quad \underline{\theta}_0 \in \Theta_0 \\ H_1 &: p(\underline{R}|\underline{\theta}_1, H_1) \quad , \quad \underline{\theta}_1 \in \Theta_1 \, . \end{aligned} \tag{2.1}$$

Based upon the observation vector, the processor must make a decision (D_0 or D_1) as to which hypothesis it believes is true. Classical detection theory has shown that decisions based upon the likelihood ratio are optimum for a wide range of goodness criteria (Peterson, Birdsall, and Fox, 1954 ; Middleton and Van Meter, 1954)

$$\Lambda(\underline{R}) \triangleq \frac{p(\underline{R}|H_1)}{p(\underline{R}|H_0)} \underset{D_0}{\overset{D_1}{\gtrless}} \eta \tag{2.2}$$

Birdsall has shown more generally that any optimality criterion based on "detection probability" $P(D_1|H_1)$ and "false alarm probability" $P(D_1|H_0)$ where good decisions are preferred over bad leads to the calculation of $\Lambda(\underline{R})$ as the decision statistic (Birdsall, 1973). Thus, a separation is achieved between the processing of $\underline{R}$ and the actual optimality criterion chosen which arises in the selection of a threshold value η.

The situation may arise where one or several of the conditioning parameters in either or both $\underline{\theta}_0$ and $\underline{\theta}_1$ are uncertain. These are then modeled as random variables and any prior knowledge about them is summarized by a priori probability density functions $p(\underline{\theta}_0)$ and $p(\underline{\theta}_1)$. The desired decision statistic now becomes the likelihood ratio of marginal probability density functions on $\mathcal{R}$

$$\Lambda(\underline{R}) = \frac{\int_{\theta_1} p(\underline{R}|\underline{\theta}_1, H_1) p(\underline{\theta}_1) d\underline{\theta}_1}{\int_{\theta_0} p(\underline{R}|\underline{\theta}_0, H_0) p(\underline{\theta}_0) d\underline{\theta}_0} . \tag{2.3}$$

Processor Structure

Once the mathematics of the likelihood ratio has been written, any realization of $\Lambda(\underline{R})$ will achieve identical results. None the less, structuring the processor in various ways often can be advantageous from the standpoint of any potential insight gained, comparison with non-Bayesian approaches to a similar problem, or the desire of greater feasibility and flexibility of implementation. Already considered, the one shot processor simply calculates $\Lambda(\underline{R})$ as shown in (2.3). Three other specific structures to be discussed are: (1) pseudo estimator, (2) sequential, and (3) two step.

The pseudo estimator structure is actually a particular case of a more general class of structures resulting from the application of Bayes' rule

$$p(\underline{\theta}|\underline{R}) = \frac{p(\underline{R}|\underline{\theta}) p(\underline{\theta})}{p(\underline{R})} . \tag{2.4}$$

Here the a posteriori probability density function of θ is calculated based on the observation vector $\underline{R}$ and the prior knowledge $p(\underline{\theta})$. The marginal density appearing in the denominator of (2.4) is simply a normalizing constant

$$p(\underline{R}) = \int_{\theta} p(\underline{R}|\underline{\theta}') p(\underline{\theta}') d\underline{\theta}' . \tag{2.5}$$

Utilizing Bayes' rule in (2.3), the likelihood ratio may be written as

$$\Lambda(\underline{R}) = \frac{p(\underline{\theta}_1) p(\underline{\theta}_0|\underline{R}, H_0)}{p(\underline{\theta}_0) p(\underline{\theta}_1|\underline{R}, H_1)} \quad \Lambda(\underline{R}|\underline{\theta}_1, \underline{\theta}_0) \tag{2.6}$$

where $$\Lambda(\underline{R}|\underline{\theta}_1,\underline{\theta}_0) = \frac{p(\underline{R}|\underline{\theta}_1,H_1)}{p(\underline{R}|\underline{\theta}_0,H_0)}$$

i.e., the parameters known likelihood ratio. Note that any convenient, admissible value of the parameters may be used to evaluate (2.6).

A pseudo estimator structure results if values for $\underline{\theta}_0$ and $\underline{\theta}_1$ may be found such that

$$\frac{p(\underline{\theta}_1)p(\underline{\theta}_0|\underline{R},H_0)}{p(\underline{\theta}_0)p(\underline{\theta}_1|\underline{R},H_1)} = 1 \ . \tag{2.7}$$

Solution values thus chosen for $\underline{\theta}_0$ and $\underline{\theta}_1$ are called pseudo estimates (Jaarsma, 1969). One value of such a structure is for the purpose of comparison with other ad hoc processors where some estimation scheme is joined with the parameters known likelihood ratio to yield a suboptimal, but perhaps easily implementable, receiver design (Hatsell and Nolte, 1974).

A second approach to processing the observation vector is to do so sequentially, i.e., the processor operates on one or a small block of observations at a time in a serial fashion until all the data has been exhausted. Consider a vector of dimension K

$$p(\underline{R}_K) = \prod_{i=1}^{K} p(R_i|\underline{R}_{i-1}) \ . \tag{2.8}$$

An expression for $p(R_i|\underline{R}_{i-1})$ is desired. Assuming parameter conditional independence

$$\begin{aligned} p(R_i|\underline{R}_{i-1}) &= \int_\theta p(R_i|\underline{R}_{i-1},\underline{\theta})p(\underline{\theta}|\underline{R}_{i-1})d\underline{\theta} \\ &= \int_\theta p(R_i|\underline{\theta})p(\underline{\theta}|\underline{R}_{i-1})d\underline{\theta} \end{aligned} \tag{2.9}$$

where $p(\underline{\theta}|\underline{R}_{i-1})$ is an updated version of the a priori probability density function of $\underline{\theta}$

$$p(\underline{\theta}|\underline{R}_{i-1}) = \frac{p(\underline{R}_{i-1}|\underline{\theta})p(\underline{\theta}|\underline{R}_{i-2})}{p(\underline{R}_{i-1}|\underline{R}_{i-2})} \quad . \tag{2.10}$$

The expressions (2.8), (2.9), and (2.10) when conditioned on H_1 and H_0 are the design equations used to obtain the marginal distributions in (2.2) for K iterations. In general, the numerator and denominator equations in (2.2) must remain separated in the updating sequence.

Principal advantages gained in structuring a processor sequentially include no need to specify the actual total observation length and the inherent learning or adaptive nature of the processor through the iterative updating of its a priori knowledge of the uncertain parameters. Furthermore, Nolte has shown that it may be necessary to implement some processors sequentially in order to avoid feasibility problems such as a growing memory requirement (Nolte, 1965; Nolte, 1966)

The final processing structure to be discussed concerns a two step approach proposed by Birdsall (Birdsall, 1968; Birdsall and Gobien, 1973). In the primary processor, the observation vector is processed in conjunction with any convenient densities (subject to minor restrictions) substituted for the actual a priori knowledge. The output of the primary processor is utilized by the secondary processor along with the true a priori knowledge to calculate the likelihood ratio. A more detailed description will require introducing the concepts of sufficient statistics and reproducing densities. These concepts will be most important in the development of later chapters in this dissertation.

From a Bayesian point of view, a fixed finite dimensional sufficient statistic of the observation for an unknown parameter vector would be defined as

$$p(\underline{\theta}|\underline{R}_K) = p(\underline{\theta}|\underline{\delta}(\underline{R}_K)) \quad , \quad \text{for all } \underline{R}_K. \tag{2.11}$$

Thus, the statistic $\underline{\delta}(\underline{R}_K)$ contains as much information about $\underline{\theta}$ as do the observations themselves. Furthermore, the dimension of $\underline{\delta}(\underline{R}_K)$ remains constant even as the dimensionality of the observation vector increases.

The classical factorization theorem provides conditions for identifying a sufficient statistic.

Theorem 2.1 Let $p(\underline{R}_K|\underline{\theta})$ be the conditional density and $\underline{\delta}(\underline{R}_K)$ the fixed finite dimensional statistic as defined previously. Then $\underline{\delta}(\underline{R}_K)$ is sufficient for $\underline{\theta}$ if there exist:

(1) a function $g[\underline{\delta}(\underline{R}_K),\underline{\theta}]$ which depends on the observation only through $\underline{\delta}(\cdot)$, and

(2) a function $G(\underline{R}_K)$ which does not depend on $\underline{\theta}$, such that

$$p(\underline{R}_K|\underline{\theta}) = g[\underline{\delta}(\underline{R}_K),\underline{\theta}]\, G(\underline{R}_K)\,. \tag{2.12}$$

Consider the following example to illustrate this concept. The observations R_i consist of an unknown scalar θ added to independent samples n_i drawn from a distribution $N(0,\sigma^2)$.

$$R_i = \theta + n_i \qquad i = 1,2,\ldots$$

Since the observations are parameter conditionally independent, their joint distribution may be written and factored as follows

$$\begin{aligned} p(\underline{R}_K|\theta) &= \prod_{i=1}^{K} p(R_i|\theta) \\ &= (2\pi\sigma^2)^{-K/2} \exp[-\frac{1}{2\sigma^2}\sum_{i=1}^{K} R_i^2] \cdot \exp[-\frac{1}{2\sigma^2}(K\theta^2 - 2\theta\sum_{i=1}^{K} R_i)] \\ &= G(\underline{R}_K)\; g[\underline{\delta}(\underline{R}_K),\theta] \end{aligned} \tag{2.13}$$

where $\underline{\delta}(\underline{R}_K) = [K, \sum_{i=1}^{K} R_i]^T$.

By simply applying Bayes' rule, we can show that if $\underline{\delta}(\underline{R}_K)$ is sufficient for $\underline{\theta}$, then the a posteriori density of $\underline{\theta}$ given $\underline{\delta}(\underline{R}_K)$ is independent of $\underline{R}_K$. Substituting (2.12) into (2.4) we see that $G(\underline{R}_K)$ cancels between the numerator and denominator leaving

$$p(\underline{\theta}|\underline{R}_K) = \frac{g[\underline{\delta}(\underline{R}_K),\underline{\theta}]p(\underline{\theta})}{\int_\Theta \text{(numerator)}\ d\underline{\theta}} . \tag{2.14}$$

Next, we wish to consider a definition and theorem on reproducing density functions (Birdsall and Gobien, 1973).

Definition 2.1 Let $\mathcal{H}_\Gamma(\theta) = \{h(\underline{\theta};\underline{\gamma})\ ;\ \underline{\gamma}\in\Gamma\subset R^m,\ \underline{\theta}\in\Theta\}$ be a family of pdf's on θ which is indexed by the m-dimensional parameter $\underline{\gamma}$. $\mathcal{H}_\Gamma(\theta)$ is said to be a reproducing class of probability densities under $p(\underline{R}_K|\underline{\theta})$ if, for any K, whenever the a priori pdf on θ is

$$p'(\underline{\theta}) = h(\underline{\theta};\underline{\gamma}^0) \quad , \quad \underline{\gamma}^0 \in \Gamma$$

there exists a $\underline{\gamma}^K = \underline{\gamma}^K(\underline{\gamma}^0,\underline{R}_K)\in\Gamma$ such that the a posteriori pdf is

$$p'(\underline{\theta}|\underline{R}_K) = h(\underline{\theta},\underline{\gamma}^K) \quad , \quad \underline{\gamma}^K \in \Gamma .$$

Thus, the a posteriori pdf remains in the same family of functions as the a priori pdf, differing only in the values of the parameters characterizing members in the family. Primes will be used to signify that we are within the class of natural reproducing densities.

Theorem 2.2 Suppose $p(\underline{R}_K|\underline{\theta})$ admits a sufficient statistic of fixed dimension $\underline{\delta}(\underline{R}_K)$ for $\underline{\theta}$ and hence can be factored as in (2.12); let the function $g(\cdot,\cdot)$ be as defined there, and, provided the integral exists, put

$$p'(\underline{\theta};\underline{\gamma}) = \frac{g[\underline{\gamma},\underline{\theta}]}{\int_\Theta g[\underline{\gamma},\underline{\theta}']d\underline{\theta}'} \quad , \quad \underline{\gamma}\in\Gamma \tag{2.15}$$

where Γ is the image of the space of observations under $\underline{\delta}(\cdot)$. Then $\{p'(\underline{\theta};\underline{\gamma}) \; ; \; \underline{\gamma} \in \Gamma\}$ is a reproducing class of densities under $p(\underline{R}_K|\underline{\theta})$. The class thus defined is called the natural conjugate class of pdf's under $p(\underline{R}_K|\underline{\theta})$; existence of a sufficient statistic implies existence of such a class.

It is possible that the natural conjugate class may not contain a member suitable for describing the true a priori knowledge. However, suppose the true a priori pdf on Θ can, for some $\underline{\gamma}^0 \in \Gamma$, be written

$$p(\underline{\theta}) = r(\underline{\theta})p'(\underline{\theta};\underline{\gamma}^0) \tag{2.16}$$

where $r(\underline{\theta})$ is a nonnegative function defined on Θ and $p(\underline{\theta})$ is absolutely continuous with respect to $p'(\underline{\theta};\underline{\gamma}^0)$. Since $p'(\underline{\theta};\underline{\gamma}^0)$ is reproducing, a simple application of Bayes' rule (2.4) reveals that $p(\underline{\theta})$ also reproduces with parameter $\underline{\gamma}$

$$p(\underline{\theta}|\underline{R}_K) = \frac{r(\underline{\theta})p'(\underline{\theta};\underline{\gamma}^K)}{\int_\Theta \text{(numerator)} \; d\underline{\theta}} . \tag{2.17}$$

Utilizing (2.17) and a Bayes' rule substitution for $p'(\underline{\theta};\underline{\gamma}^K)$, the marginal distribution of $\underline{R}_K$ given the true a priori knowledge can be written in terms of the natural conjugate class of densities and $r(\underline{\theta})$

$$p(\underline{R}_K) = p'(\underline{R}_K) \int_\Theta r(\underline{\theta})p'(\underline{\theta};\underline{\gamma}^K) \, d\underline{\theta} . \tag{2.18}$$

The mathematical description of the two step approach to processor structure now can be completed. The primary processor uses a convenient description of a priori knowledge out of the class of natural conjugate densities (if such a class exists). The secondary processor utilizes the likelihood ratio calculated on the basis of these densities together with the resulting sufficient statistics characterizing the a posteriori pdf's and the true a priori knowledge to calculate the true likelihood ratio

$$\Lambda(\underline{R}_K) = \Lambda'(\underline{R}_K) \frac{\int_{\Theta_1} r(\underline{\theta}_1) p'(\underline{\theta}_1; \underline{\gamma}_1^K)\, d\underline{\theta}_1}{\int_{\Theta_0} r(\underline{\theta}_0) p'(\underline{\theta}_0; \underline{\gamma}_0^K)\, d\underline{\theta}_0} \tag{2.19}$$

where $$\Lambda'(\underline{R}_K) = \frac{\int_{\Theta_1} p(\underline{R}_K|\underline{\theta}_1) p'(\underline{\theta}_1)\, d\theta_1}{\int_{\Theta_0} p(\underline{R}_K|\underline{\theta}_0) p'(\underline{\theta}_0)\, d\underline{\theta}_0}.$$

The benefit of such an approach to receiver design is that potentially a major portion of the processor can be designed without knowing the exact a priori knowledge. Furthermore, the mathematical tractability of the natural conjugate priors may simplify the design of the primary processor.

Performance

The complete description of a detection device includes both the processor itself (i.e., the mathematical transformation from observation space to decision statistic) and the performance of the processor evaluated with respect to the goodness criterion initially chosen. As mentioned earlier, the likelihood ratio has been shown optimum for any goodness criterion based on "detection probability" $P(D_1|H_1)$ and "false alarm probability" $P(D_1|H_0)$ where good decisions are preferred over bad. Thus, the appropriate description of performance for a likelihood ratio computing device is its detection and false alarm probabilities as a function of decision threshold. The precise definition of these terms (which arise from within a RADAR and SONAR context) now will be given.

Since the likelihood ratio is simply a transformation of random variables (the observation vector $\underline{R}$) to a one dimensional decision statistic ($\Lambda(\underline{R})$), the likelihood ratio itself will be a random variable whose probability density function will depend on which hypothesis (H_0 or H_1)

is actually active on $\mathcal{R}$. Recalling from (2.2) that the threshold η divides the decision space, define

$$P_D \triangleq p(D_1|H_1) \triangleq \int_\eta^\infty p(\Lambda|H_1)\, d\Lambda \tag{2.20}$$

$$P_F \triangleq p(D_1|H_0) \triangleq \int_\eta^\infty p(\Lambda|H_0)\, d\Lambda \ . \tag{2.21}$$

Peterson, Birdsall, and Fox introduced a graphical representation of P_D versus P_F as a function of η known as the ROC (receiver operating characteristic) curve (Peterson, Birdsall, and Fox, 1954). The ROC curve will be the means by which performance of the detection receivers discussed in this dissertation will be evaluated and compared.

In general, the entire ROC curve is necessary to completely specify performance. However, in the classic SKE in WGN problem (H_1: signal known exactly + white Gaussian noise vs. H_0: white Gaussian noise alone), performance is summarized by a single number known as the detectability index d^2. In this case, the distribution of $\ell(\underline{R}) = \ln\Lambda(\underline{R})$ is Gaussian under H_1 and H_0 with equal variances of $2E/N_0$ and means separated by $2E/N_0$ (E = received signal energy; $N_0/2$ = noise power spectrum height). By definition (Van Trees, 1968)

$$d^2 \triangleq \frac{[E(\ell|H_1) - E(\ell|H_0)]^2}{\mathrm{var}(\ell|H_0)} \tag{2.22}$$

$$= \frac{2E}{N_0}$$

Detection and false alarm probability expressions corresponding to (2.20) and (2.21) are

$$P_D = \mathrm{erfc}_*\left(\frac{\ln \eta}{d} - \frac{d}{2}\right) \tag{2.23}$$

$$P_F = \text{erfc}_*\left(\frac{\ln \eta}{d} + \frac{d}{2}\right) \tag{2.24}$$

where

$$\text{erfc}_* = \int_x^\infty \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{x^2}{2}\right) dx \ . \tag{2.25}$$

The SKE performance curves are illustrated in Figure 2.1 on normal-normal paper. Note that performance increases linearly on the negative diagonal as a function of d.

Block Diagrams and Sufficient Statistics

From the preceding discussion, it has been shown that once the likelihood ratio was determined, $\Lambda(\underline{R})$ could be implemented in various structures which might look quite different. Perhaps the epitome of this is in Birdsall's two step approach where the primary processor might take on any one of an infinite number of structures depending on the tractable prior chosen. Nevertheless, it is quite common for those of us engaged in optimum receiver design to look for structural pieces in our particular realization of $\Lambda(\underline{R})$ that might be either the same as, or in contrast to, structure arrived at by other ad hoc approaches to the same problem. The contention here is that perhaps too much emphasis has been placed in the past on the overall structure of $\Lambda(\underline{R})$ instead of some more fundamental component pieces. A solution for the likelihood ratio not in closed form always has appeared only half completed.

The expression in (2.3) shows that the fundamental components of $\Lambda(\underline{R})$ before the introduction of a particular a priori knowledge are the conditional densities of $\underline{R}$ under H_1 and H_0. Assuming $p(\underline{R}|\underline{\theta})$ admits a sufficient statistic for $\underline{\theta}$, (2.12) shows that the two basic building blocks are the sufficient statistic $\underline{\delta}(\underline{R})$ and $G(\underline{R})$. In general, no matter what a priori knowledge is chosen, $\underline{\delta}(\underline{R})$ and $G(\underline{R})$ will have to be calculated

under H_1 and H_0. Any final structure of $\Lambda(\underline{R})$ will always be a function of these basic parts and it is their structure which will provide the most fundamental basis from which comparison with other detectors can be made. As an example, a recent paper by Adams and Nolte derives and discusses the interpretation of these basic components for optimum array processors in fluctuating ambient noise fields (Adams and Nolte, 1975).

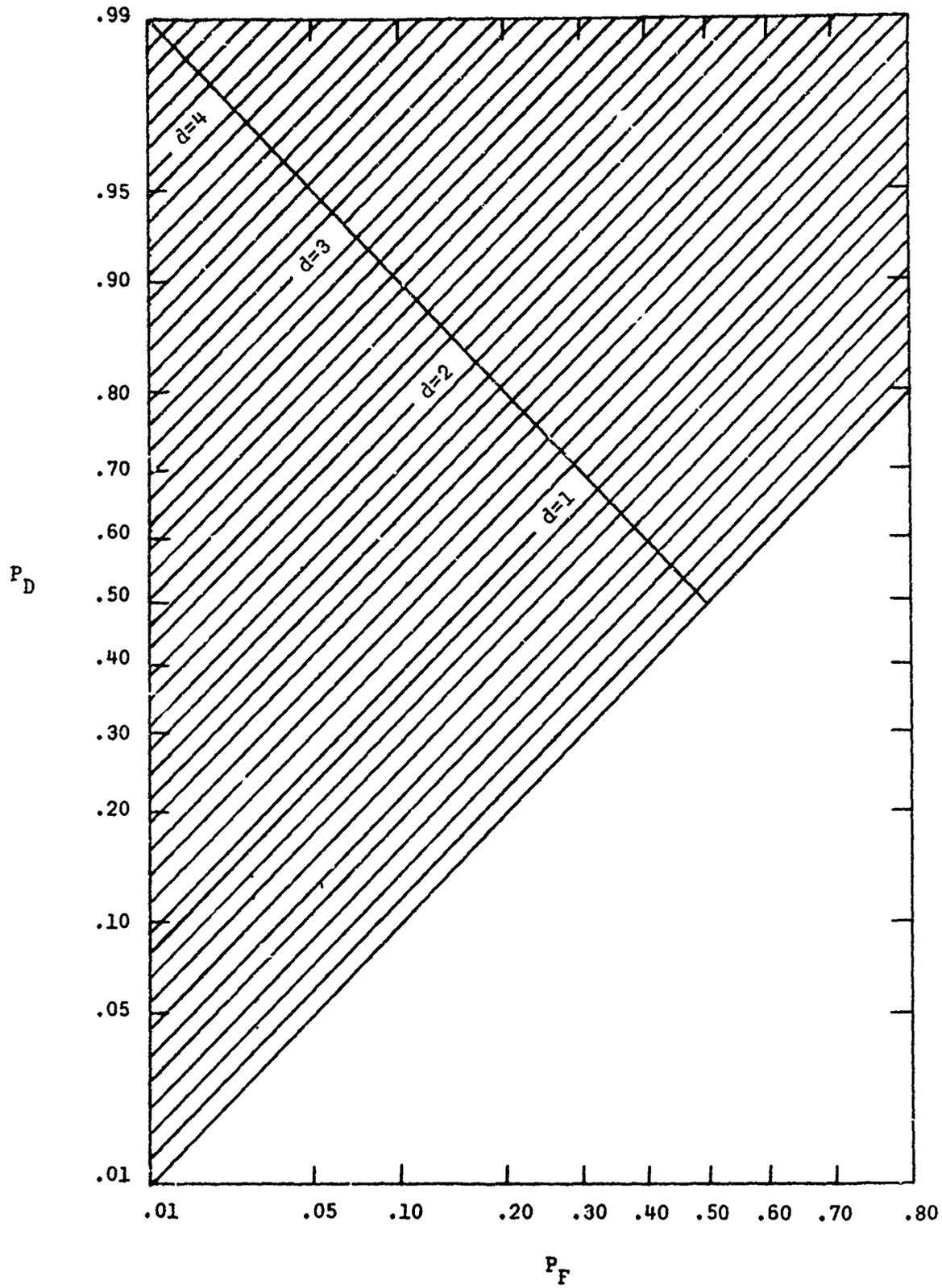


Figure 2.1. Performance of the SKE Processor.

Chapter III

OPTIMUM ARRAY PROCESSING

A rigorous approach to signal detection theory from a Bayesian point of view has been established. The desire now is to apply these general concepts to the optimal processing of data from an array of sensors. It should be emphasized that the Bayesian approach is global in the sense that no processing structure (such as a beam former) is assumed from the beginning. Rather, the mathematics (which results in structure) is allowed to evolve freely out of the problem statement with the unprocessed outputs of each array element taken as observables. Furthermore, since the likelihood ratio will be used as the decision statistic, the only complete description of array processor performance will be in the form of an ROC curve.

The particular interest in this dissertation will be the general formulation of sequential array processors when there is directional uncertainty. Cases will be considered where there is both a far field directional signal source and noise source in addition to an independent additive noise component at each array element.

Literature Review

The Bayesian approach to array processing has been discussed in the literature since the early 1960's. Unfortunately, there have been few instances where the likelihood ratio processor's performance has been given in terms of the ROC curve. Quite often, other performance measures are used (such as array gain) which are not the optimality criterion under

which the processor was designed. A brief summary of the literature related to this dissertation is given below. An excellent topical and chronological review of optimum array processing can be found in a recent dissertation by Adams (Adams, 1973) and paper by Adams and Nolte (Adams and Nolte, 1975).

The earliest attempts to derive Bayes optimal array processors assumed known directional signal sources and noise fields which at times contained a highly directional additive component of known location. Bryn was the first with his paper in 1962 on the detection of a Gaussian signal in Gaussian noise (Bryn, 1962). His formulation used a truncated power series expansion for the likelihood ratio which was optimum at low signal-to-noise ratios. Unfortunately, Bryn's evaluation of performance was in terms of array gain instead of detectability. A short time later, Middleton and Groginsky discussed in general terms the same problem without making a low signal-to-noise ratio assumption (Middleton and Groginsky, 1964). Their study was mainly concerned with factorization results (i.e., the splitting of processor structure into a spatial component independent of noise statistics and a component dependent only on the statistics of the signal and noise random processes) and no performance was given. Schultheiss also considered the problem of detecting a Gaussian signal in Gaussian noise (Schultheiss, 1968). His analysis centered on the benefit of likelihood ratio processing over conventional beam forming when a major component of the noise field was highly directional. True detection performance was calculated based on a low signal-to-noise ratio assumption.

Another paper which quickly followed Bryn's was the first to deal with detecting a signal of known form in Gaussian noise (Stocklin, 1963). Stocklin's work is to be noted for its complete analysis from likelihood

ratio through ROC curve. In 1964, Mermoz discussed the same problem from a different point of view (Mermoz, 1964). His approach was to postulate a linear structure consisting of filters at each array element whose outputs were then summed together. The filter transfer functions were calculated which maximized the signal-to-noise ratio at the summing junction. For this particular problem, the resulting processor is identical to that of the Bayes optimal array processor. As in Middleton's and Groginsky's work, a major contribution of this paper was in its factorization results. Mermoz's analysis also is presented in a book by Horton (Horton, 1969).

Several years later, the solutions due to both Bryn and Mermoz were rederived by Cox from within a common mathematical framework using the Schwartz inequality to maximize the detectability index (Cox, 1968; Cox, 1969). Particularly well presented is the concept that such seemingly diverse problems are actually closely related. In addition, the 1968 paper gives an excellent discussion of detection theory results in terms of complex random vectors with emphasis on the trigonometric Fourier series.

The incorporation of directional uncertainty into Bayes optimal array processors has occurred only recently. Young and Howard modeled their signal as having a uniform a priori location distribution over a particular field of interest (Young and Howard, 1970). Performance was given in terms of average risk. In addition, average risk for the optimal processor was compared to that of the conventional beam former as a demonstration of the latter's rather severe sensitivity to location uncertainty. The next significant step was taken by Gallop and Nolte (Gallop, 1971; Gallop and Nolte, 19[illegible] Their analysis allowed signal location uncertainty to be modeled as lying on a continuum of a priori densities, extremes of

which were uniform knowledge on one end and certainty of location on the other. Sets of ROC curves representing various levels of location uncertainty were given.

A general approach to the derivation of array processors for sources of known location was proposed in a paper by Adams and Nolte (Adams and Nolte, 1975). Particular emphasis was given to receiver structure, factorization results, and performance when the noise field contained an additive component from a fixed far field location. Additionally, Adams also has pursued the derivation of likelihood ratio expressions where either the signal or directional noise locations were uncertain (Adams, 1973). No performance for these cases was presented. Essentially, the work in this dissertation is an extension of that by Adams to the consideration of the sequential implementation of array processors when there is directional uncertainty and the calculation of performance for several specific problems.

Trigonometric Fourier Series

The application of classical statistics to optimal detection theory requires a finite dimensional vector of observables whose joint probability density function under H_0 and H_1 is used to form the likelihood ratio. Typically, the array element outputs will be a collection of continuous time waveforms. Some mapping then must be used to carry this space into into a space of finite dimensional vectors. Three well known approaches to this transformation are: (1) time sampling, (2) the Karhunen-Loève expansion, and (3) the trigonometric Fourier series. The selection of a particular mapping is heavily problem dependent. A wise choice can often greatly simplify the mathematics of the likelihood ratio. Adams

showed that the third approach led to a significant amount of tractability, especially when the noise field contained an additive source in the far field (Adams, 1973; Adams and Nolte, 1975). The trigonometric Fourier series will be used throughout this dissertation. Basic definitions and results for the characterization of random processes follow (see also Papoulis, 1965, and Van Trees, 1968).

What we desire is a series expansion valid for sample functions z(t) from a zero mean stationary random process over the interval (-T/2,T/2). Written as a Fourier series

$$z(t) = \underset{N\to\infty}{\text{l.i.m.}} \sum_{n=-N}^{N} z(n) \left(\frac{1}{T}\right)^{1/2} \exp(jn\omega_0 t) \ , \ |t|<T/2 \qquad (3.1)$$

where $$\omega_0 = \frac{2\pi}{T}$$

and $$z(n) = \left(\frac{1}{T}\right)^{1/2} \int_{-T/2}^{T/2} z(t) \exp(-jn\omega_0 t)\, dt \ . \qquad (3.2)$$

Convergence is in the mean square sense (Papoulis, 1965). The notation "l.i.m." denotes limit in the mean which is defined as

$$\lim_{N\to\infty} E\left[\left(z(t) - \sum_{n=-N}^{N} z(n) \left(\frac{1}{T}\right)^{1/2} \exp(jn\omega_0 t)\right)^2\right] = 0 \ , \ |t|<T/2 \ . \qquad (3.3)$$

Since z(t) is a random process, z(n) will be a random variable and (Papoulis, 1965)

$$\lim_{T\to\infty} E[z(n)z(m)^*] = \begin{cases} N(n\omega_0) & , \ n = m \\ 0 & , \ n \neq m \end{cases} \qquad (3.4)$$

where $N(n\omega_0)$ is the power spectral density of the random process and "*" denotes conjugate transpose. Thus, as the observation interval T increases, the Fourier coefficients at different frequency indices become uncorrelated.

The series in (3.1) can be truncated at any N with a corresponding representation error. The expression in (3.3) guarantees that the expected value of this error becomes smaller as N is increased. Particularly with bandlimited random processes, (3.4) indicates that the expected value of error will be negligible when $N = 2\pi W/\omega_0$ (W the bandwidth in Hz.) for T taken sufficiently large. Under these conditions, we will say that the real sample function z(t) can be mapped into the finite dimensional vector $\underline{z}$ where

$$\underline{z} = [z(0),\ldots,z(N)]^T \tag{3.5}$$

the "T" denoting transpose. Coefficients with negative indices are not written since for real functions $z(-n) = z(n)^*$. Considering such a finite dimensional representation, if z(t) is replaced by its series expansion as in (3.1), then the energy of the sample function over the observation interval (-T/2,T/2) can be approximated by

$$\int_{-T/2}^{T/2} z(t) \cdot z(t)\, dt \approx \sum_{n=-N}^{N} z(n)^* z(n) . \tag{3.6}$$

From (3.4), we see that for T sufficiently large

$$E\left[\int_{-T/2}^{T/2} z(t) \cdot z(t)\, dt\right] \approx \sum_{n=-N}^{N} N(n\omega_0) . \tag{3.7}$$

The discussion so far has emphasized random processes. Convergence in the mean square sense guaranteed that almost every sample function could be represented by its series expansion. More strict conditions can be applied in the case of a deterministic time waveform s(t). Since $\{(1/T)^{1/2} \exp(jn\omega_0 t)\ ;\ n = -\infty,\ldots,\infty\}$ forms a complete orthonormal set, then for all s(t) with finite energy in the interval (-T/2,T/2)

$$s(t) = \lim_{N\to\infty} \sum_{-N}^{N} s(n) \left(\frac{1}{T}\right)^{1/2} \exp(jn\omega_0 t) \quad , \quad |t|<T/2 \qquad (3.8)$$

where $$\omega_0 = \frac{2\pi}{T}$$

and $$s(n) = \left(\frac{1}{T}\right)^{1/2} \int_{-T/2}^{T/2} s(t) \exp(-jn\omega_0 t) \, . \qquad (3.9)$$

Again, convergence is in the mean square sense with (Van Trees, 1968)

$$\lim_{N\to\infty} \int_{-T/2}^{T/2} [s(t) - \sum_{n=-N}^{N} s(n) \left(\frac{1}{T}\right)^{1/2} \exp(jn\omega_0 t)]^2 \, dt = 0. \qquad (3.10)$$

Furthermore, the energy of s(t) over the observation interval (-T/2,T/2) is

$$\int_{-T/2}^{T/2} s(t) \cdot s(t) \, dt = \sum_{n=-\infty}^{\infty} s(n)^* \, s(n) \qquad (3.11)$$

which is simply Parseval's theorem.

The optimal array processors to be considered in this dissertation observe a vector of real time waveforms on the interval (-T/2,T/2)

$$\underline{z}(t) = [z_0(t),\ldots,z_{K-1}(t)]^T \qquad (3.12)$$

where the subscript denotes the array element. Making the assumptions that led to (3.5), let

$$z_k(n) = \left(\frac{1}{T}\right)^{1/2} \int_{-T/2}^{T/2} z_k(t) \exp(-jn\omega_0 t) \, dt \qquad (3.13)$$

$$\underline{z}(n) = [z_0(n),\ldots,z_{K-1}(n)]^T \qquad (3.14)$$

and $$\underline{z} = [\underline{z}(0)^T,\ldots,\underline{z}(N)^T]^T \, . \qquad (3.15)$$

In this way, the time waveforms observed on the K elements are mapped into a K·(N+1) dimensional vector. Note that the Fourier coefficients for a single frequency index n and all K elements are grouped together.

Multivariate Complex Gaussian Distributions

The particular arrangement of Fourier coefficients in the vector $\underline{z}$ will lead to certain mathematical simplifications in the likelihood ratio calculations to be discussed later. In preparation, this section will develop the necessary theory of multivariate complex Gaussian distributions. The presentation closely follows that of Adams and Nolte (Adams and Nolte, 1975) which is based on Goodman (Goodman, 1963).

First, we will need definitions for complex Gaussian random variables and vectors. A complex Gaussian random variable is defined to be a complex random variable whose real and imaginary parts are bivariate Gaussian. A complex Gaussian random vector of dimension p is defined to be a p-tuple of complex Gaussian random variables such that the vector of real and imaginary parts is 2p-variate Gaussian.

Consider the complex Gaussian random vector $\underline{z}$ which can be written

$$\underline{z} = \underline{x} + i\underline{y} . \tag{3.16}$$

Our attention will be restricted to complex Gaussian random vectors whose covariance matricies are of the following special form. Both $\underline{x}$ and $\underline{y}$ have covariance matricies which are equal to 1/2 $\underline{V}$, where $\underline{V}$ is a symmetric, positive semi-definite matrix. The covariance of $\underline{x}$ and $\underline{y}$ is assumed equal to 1/2 $\underline{W}$, where $\underline{W}$ is skew symmetric and

$$\begin{aligned} E\ [\underline{x} - E(\underline{x})][\underline{y} - E(\underline{y})]^T &= \frac{1}{2}\underline{W} \\ E\ [\underline{y} - E(\underline{y})][\underline{x} - E(\underline{x})]^T &= -\frac{1}{2}\underline{W} . \end{aligned} \tag{3.17}$$

Under these conditions, the p-variate complex Gaussian distribution of $\underline{z}$ is given by

$$p(\underline{z}) = \frac{1}{\pi^P |\underline{Q}|} \exp[-(\underline{z}-\underline{m})^* \underline{Q}^{-1} (\underline{z}-\underline{m})] \tag{3.18}$$

where
$$\underline{m} = E\{\underline{z}\} \tag{3.19}$$

and
$$\underline{Q} = E\{(\underline{z}-\underline{m})(\underline{z}-\underline{m})^*\} = \underline{V} + i\underline{W}. \tag{3.20}$$

The Likelihood Ratio and Arry Processor Structure

The results of Chapter II will now be rewritten in the context of optimal array processors. The real random processes observed at the output of each array element will be assumed stationary and jointly Gaussian. Since (3.13) indicates that the Fourier coefficients at a particular element are simply linear functionals of a Gaussian random process, the real and imaginary parts of each $z_k(n)$ will be bivariate Gaussian distributed (Van Trees, 1968). Thus, $z_k(n)$ is a complex Gaussian random variable. Furthermore, since the observed random processes are jointly Gaussian, the collection of real and imaginary parts of all the Fourier coefficients will be jointly Gaussian (Van Trees, 1968). Thus, $\underline{z}$ is a complex Gaussian random vector of dimension $K\cdot(N+1)$.

Consider the array detection problem where the observables consist of $K\cdot(N+1)$ Fourier coefficients arranged as in (3.15). The two mutually exclusive and exhaustive hypotheses will be

$$\begin{aligned} H_0 &: \underline{z} = n(\underline{\theta}_0) \\ H_1 &: \underline{z} = s(\underline{\theta}_1) + n(\underline{\theta}_1) \end{aligned} \tag{3.21}$$

where $\underline{n}(\underline{\theta}_0)$, $\underline{n}(\underline{\theta}_1)$, and $\underline{s}(\underline{\theta}_1)$ are noise and signal vectors written explicitly as functions of parameter vectors $\underline{\theta}_0 \in \Theta_0$ and $\underline{\theta}_1 \in \Theta_1$. Since $\underline{z}$ is a complex Gaussian random vector, the likelihood ratio is written as

$$\Lambda(\underline{z}) = \frac{\int_{\theta_1} \frac{1}{|\underline{Q}(\underline{\theta}_1)|} \exp[-(\underline{z}-\underline{m}_1(\underline{\theta}_1))^* \underline{Q}_1^{-1}(\underline{\theta}_1)(\underline{z}-\underline{m}_1(\underline{\theta}_1))] p(\underline{\theta}_1) d\underline{\theta}_1}{\int_{\theta_0} \frac{1}{|\underline{Q}(\underline{\theta}_0)|} \exp[-(\underline{z}-\underline{m}_0(\underline{\theta}_0))^* \underline{Q}_0^{-1}(\underline{\theta}_0)(\underline{z}-\underline{m}_0(\underline{\theta}_0))] p(\underline{\theta}_0) d\underline{\theta}_0} \tag{3.22}$$

where $\underline{m}_1(\underline{\theta}_1) = E\{\underline{z}|\underline{\theta}_1, H_1\}$; $\underline{m}_0(\underline{\theta}_0) = E\{\underline{z}|\underline{\theta}_0, H_0\}$

and $p(\underline{\theta}_1)$ and $p(\underline{\theta}_0)$ are *a priori* probability density functions.

As a specific example, $\Lambda(\underline{z})$ for the problem of detecting a signal known exactly in bandlimited white Gaussian noise is easily written. Since there are no uncertain parameters under either hypothesis and $\underline{m}_0 = \underline{0}$, (3.22) becomes

$$\Lambda(\underline{z}) = \exp[-\underline{m}^* \underline{Q}^{-1} \underline{m} + 2\ \mathrm{Re}\{\underline{z}^* \underline{Q}^{-1} \underline{m}\}] \tag{3.23}$$

where $\underline{m} = \underline{m}_1$

$$\underline{Q} = \underline{Q}_1 = \underline{Q}_0$$

and "Re" denotes the real part.

Recall that the performance of this processor is completely specified by the detectability index d^2 (see (2.22)) which has the value

$$d^2 = 2\ \underline{m}^* \underline{Q}^{-1} \underline{m}\ . \tag{3.24}$$

Note that under the condition in (3.4), the covariance between any two Fourier coefficients at different frequency indices will be zero. Thus, the covariance matrix $\underline{Q}$ will be block diagonal (KxK blocks) and its inverse also will be block diagonal. The expression in (3.23) can be rewritten as

$$\Lambda(\underline{z}) = \exp[-\sum_{n=0}^{N} \underline{m}(n)^* \underline{Q}^{-1}(n)\underline{m}(n) + 2\,\mathrm{Re}\{\sum_{n=0}^{N} \underline{z}(n)^* \underline{Q}^{-1}(n)\underline{m}(n)\}]$$

$$= \prod_{n=0}^{N} \exp[-\underline{m}(n)^* \underline{Q}^{-1}(n)\underline{m}(n) + 2\,\mathrm{Re}\{\underline{z}(n)^* \underline{Q}^{-1}(n)\underline{m}(n)\}] \qquad (3.25)$$

where $\underline{m}(n) = E\{\underline{z}(n)|H_1\}$ ($\underline{z}(n)$ as in (3.14))

and $\underline{Q}(n) = E\{[\underline{z}(n)-\underline{m}(n)][\underline{z}(n)-\underline{m}(n)]^*|H_1\}$.

The likelihood ratio has been broken into its "single frequency" components. This is the reason for arranging the Fourier coefficients in the observation vector $\underline{z}$ as shown in (3.15) and such a decomposition will be valid for all problems where the condition in (3.4) has been met.

Table 3.1 summarizes the likelihood ratio expressions for three realizations of the array processor structure: (1) one shot, (2) pseudo estimator, and (3) two step. Figures 3.1, 3.2, and 3.3 illustrate their corresponding block diagrams. Discussion of the sequential structure is postponed until suitable conditions on the length of the incremental observation interval have been established so that both the expression in (3.4) still may be considered valid and parameter conditional independence between successive observation increments is assured.

The fundamental form of the likelihood ratio is found in the one shot optimal array processor given in Table 3.1 A and illustrated in Figure 3.1. In this structure, all of the data is processed at the same time and the likelihood ratio is obtained directly. Once the mathematics of the likelihood ratio has been written, however, any realization of $\Lambda(\underline{z})$ will achieve the same performance. As mentioned in Chapter II, structuring the optimal processor in various ways often can be advantageous from the standpoint of any potential insight gained, comparisons with non-Bayesian approaches to a similar problem, or the desire of greater feasibility and flexibility of implementation.

The pseudo estimator structure given in Table 3.1 B and illustrated in Figure 3.2 provides a means of comparing the optimal array processor with a popular ad hoc detector structure. An appealing approach to an array detection problem where uncertain parameters exist is to estimate these parameters and then plug them into the conditional likelihood ratio as if they were known exactly. When "good" estimators are used, such a structure as illustrated in Figure 3.4 appears to be operating in an optimal fashion. The processors of Bryn (Bryn, 1962) or Mermoz (Mermoz, 1964; Horton, 1969) typically are assumed as implementations of the conditional likelihood ratio. Examples of such an approach are the papers by Chang and Tuteur (Chang and Tuteur, 1971) for the detection of a Gaussian signal of known direction in a Gaussian noise field of unknown statistics and Bienvenu and Vernet (Bienvenu and Vernet, 1972) discussing a similar problem where the signal is of known form instead of Gaussian. Another example is the paper by Giraudon (Giraudon, 1972) for the detection of a known form signal of known direction imbedded in a non-stationary noise field. It is not clear, however, that piecing together locally optimal techniques (i.e., "good" estimators and a solution optimal when all parameters are known) will yield global optimality when the overall goal is good detection performance. Note that the structures in Figures 3.2 and 3.4 are equivalent only when $\hat{\underline{\theta}}_1 = \underline{\theta}_1^*$ and $\hat{\underline{\theta}}_0 = \underline{\theta}_0^*$ are the pseudo estimates of $\underline{\theta}_1$ and $\underline{\theta}_0$. Recalling from (2.7), utilizing the pseudo estimates results in the lower branch of Figure 3.2 equaling unity. Thus, the optimal array processor can be realized in an estimate and plug structure. It should be noted that the pseudo estimate is generally not equal to a well known estimate.

The third realization of the likelihood ratio as a two step processor is given in Table 3.1 C and illustrated in Figure 3.3. The usefulness of this structure is that potentially a major portion of the optimal array processor can be designed without knowing the exact a priori knowledge. Thus, a certain degree of implementation flexibility is achieved. The two step approach provides the basis for the general time sequential array processor structure discussed in Chapter V.

Summary of Array Processor Structures

A. One Shot

$$\Lambda(\underline{z}) = \frac{\int_{\Theta_1} \frac{1}{|\underline{Q}(\underline{\theta}_1)|} \exp[-(\underline{z}-\underline{m}_1(\underline{\theta}_1))^* \underline{Q}_1^{-1}(\underline{\theta}_1)(\underline{z}-\underline{m}_1(\underline{\theta}_1))]p(\underline{\theta}_1)d\underline{\theta}_1}{\int_{\Theta_0} \frac{1}{|\underline{Q}(\underline{\theta}_0)|} \exp[-(\underline{z}-\underline{m}_0(\underline{\theta}_0))^* \underline{Q}_0^{-1}(\underline{\theta}_0)(\underline{z}-\underline{m}_0(\underline{\theta}_0))]p(\underline{\theta}_0)d\underline{\theta}_0}$$

where $\underline{m}_1(\underline{\theta}_1) = E\{\underline{z}|\underline{\theta}_1, H_1\}$; $\underline{m}_0(\underline{\theta}_0) = E\{\underline{z}|\underline{\theta}_0, H_0\}$

and $p(\underline{\theta}_1)$ and $p(\underline{\theta}_0)$ are *a priori* probability density functions.

B. Pseudo Estimator

$$\Lambda(\underline{z}) = \frac{p(\underline{\theta}_1)p(\underline{\theta}_0|\underline{z},H_0)}{p(\underline{\theta}_0)p(\underline{\theta}_1|\underline{z},H_1)} \quad \Lambda(\underline{z}|\underline{\theta}_1,\underline{\theta}_0)$$

where $$\Lambda(\underline{z}|\underline{\theta}_1,\underline{\theta}_0) = \frac{|\underline{Q}_0(\underline{\theta}_0)| \exp[-(\underline{z}-\underline{m}_1(\underline{\theta}_1))^* \underline{Q}_1^{-1}(\underline{\theta}_1)(\underline{z}-\underline{m}_1(\underline{\theta}_1))]}{|\underline{Q}_1(\underline{\theta}_1)| \exp[-(\underline{z}-\underline{m}_0(\underline{\theta}_0))^* \underline{Q}_0^{-1}(\underline{\theta}_0)(\underline{z}-\underline{m}_0(\underline{\theta}_0))]}$$

and $p(\underline{\theta}_0|\underline{z},H_0)$ and $p(\underline{\theta}_1|\underline{z},H_1)$ are *a posteriori* probability density functions.

C. Two Step

$$\Lambda(\underline{z}) = \Lambda'(\underline{z}) \frac{\int_{\Theta_1} r(\underline{\theta}_1)p'(\underline{\theta}_1;\underline{\gamma}_1^K)d\underline{\theta}_1}{\int_{\Theta_0} r(\underline{\theta}_0)p'(\underline{\theta}_0;\underline{\gamma}_0^K)d\underline{\theta}_0}$$

where $$p(\underline{\theta}_1) = r(\underline{\theta}_1)p'(\underline{\theta}_1;\underline{\gamma}_1^0)$$

$$p(\underline{\theta}_0) = r(\underline{\theta}_0)p'(\underline{\theta}_0;\underline{\gamma}_0^0)$$

and $p'(\underline{\theta}_1;\underline{\gamma}_1^0)$ and $p'(\underline{\theta}_0;\underline{\gamma}_0^0)$ are the natural conjugate *a priori* probability density functions used in forming $\Lambda'(z)$.

Table 3.1

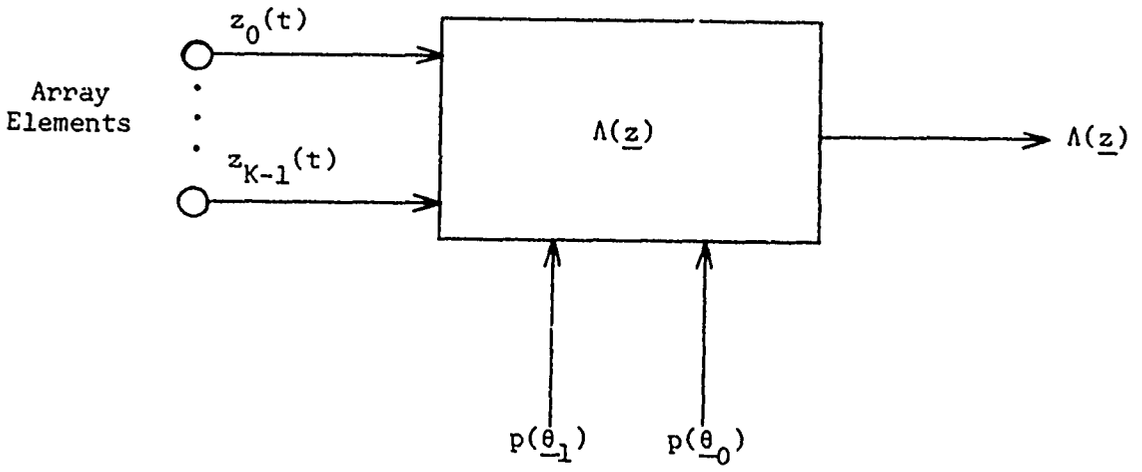


Figure 3.1. One Shot Array Processor.

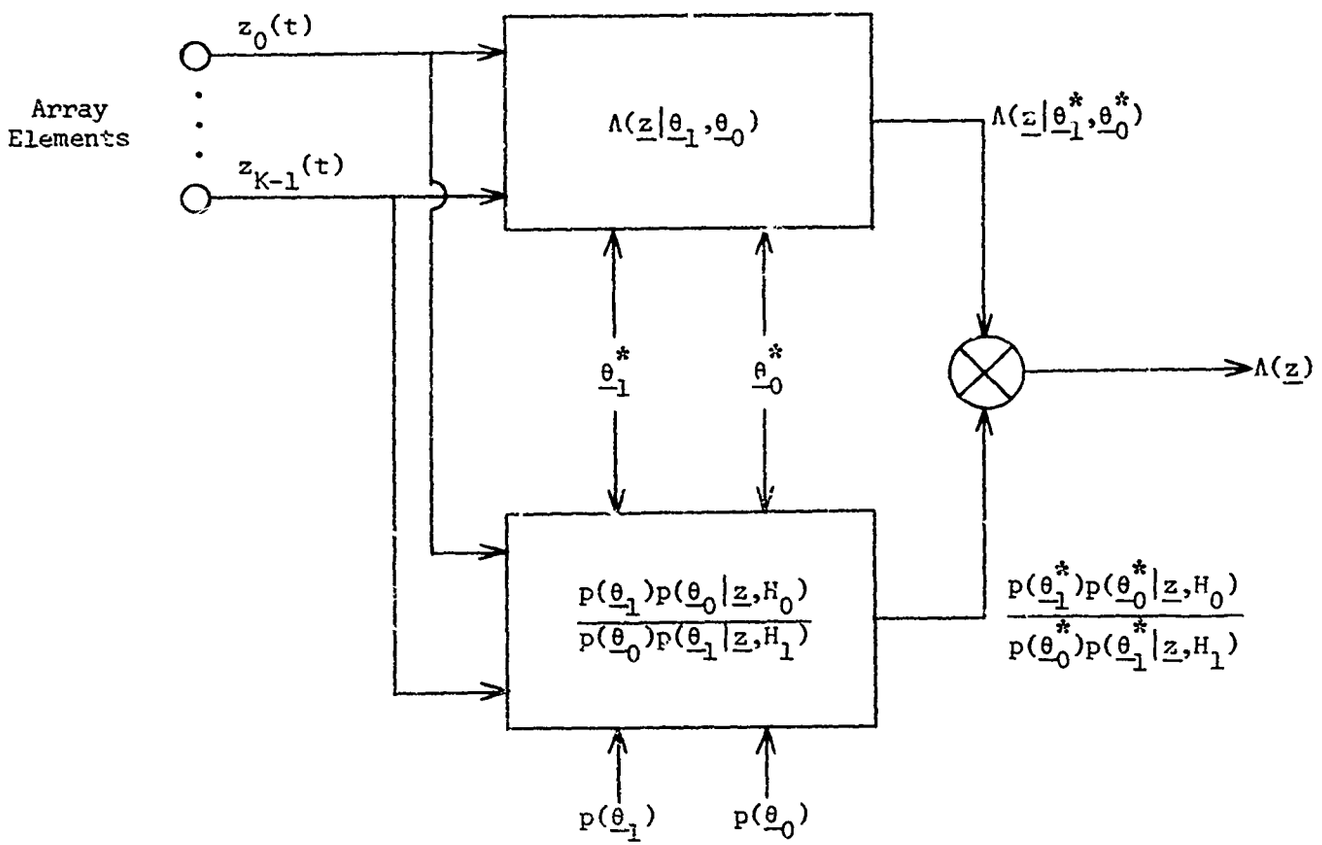


Figure 3.2. Pseudo Estimator Array Processor.

Figure 3.3. Two Step Array Processor.

arising from adjacent observation periods would be desirable since it can lead to a convenient implementation of the sequential array processor via (2.8), (2.9), and (2.10). Recall that when the observables are jointly Gaussian, zero covariance implies independence. In this chapter, equations will be derived which express the covariance between the Fourier coefficients explicitly as a function of observation period length.

Covariance Arising From a Scalar Random Process

Let z(t) be a sample function from the zero mean stationary random process observed at the output of a single array element. The Fourier coefficients for this time waveform will be as defined in Chapter III

$$z(n) = \left(\frac{1}{T}\right)^{1/2} \int_{-T/2}^{T/2} z(t)\exp(-jn\omega_0 t)dt \qquad (4.1)$$

where $\omega_0 = \frac{2\pi}{T}$.

The expression in (3.4) indicates that the Fourier coefficients at different frequency indicies become uncorrelated as the observation interval T increases. Unfortunately, little in insight is gained as to how fast this occurs. The derivation in Appendix A due to Blachman is most beneficial to this respect (Blachman, 1957). Summarizing those results

$$E[z(n)z(m)^*] = \int_{-\infty}^{\infty} N(x\omega_0)\text{sinc}(x-n)\sin(x-m)dx \qquad (4.2)$$

where $N(x\omega_0)$ = the power spectral density function of the random process

and $\text{sinc}(x) = \frac{\sin(\pi x)}{(\pi x)}$.

Note the orthonormal property of sinc(x)

$$\int_{-\infty}^{\infty} \text{sinc}(x-n)\text{sinc}(x-m)\, dx = \delta_{nm} \tag{4.3}$$

where $$\delta_{nm} = \begin{cases} 1, & n=m \\ 0, & n\neq m \end{cases} .$$

The variance and covariance of the Fourier coefficients now easily can be pictured as areas under a curve. The integrand in (4.2) is illustrated graphically in Figure 4.1 for two specific cases of the following expressions

(1) Variance

$$E[z(n)z(n)^*] = \int_{-\infty}^{\infty} N(x\omega_0)\text{sinc}^2(x-n)\, dx \tag{4.4}$$

(2) Covariance

$$E[z(n)z(m)^*] = \int_{-\infty}^{\infty} N(x\omega_0)\text{sinc}(x-n)\text{sinc}(x-m)\, dx . \tag{4.5}$$

Two conclusions may be drawn

(1) Variance

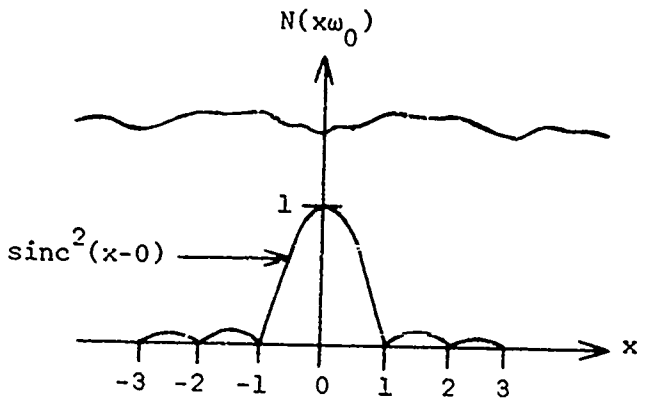
As long as $N(x\omega_0)$ is approximately constant within $2\omega_0$ or $3\omega_0$ either side of $n\omega_0$, then

$$E[z(n)z(n)^*] \simeq N(n\omega_0) . \tag{4.6}$$

(2) Covariance

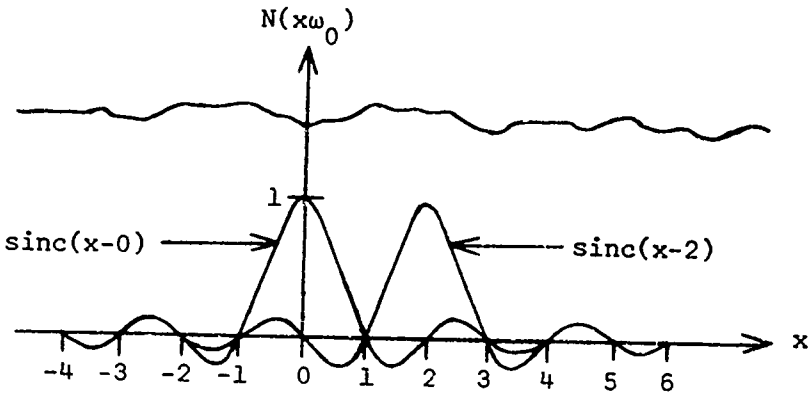
As long as $N(x\omega_0)$ is approximately constant over the interval where the product $\text{sinc}(x-n)\cdot\text{sinc}(x-m)$ has appreciable value, then

$$E[z(n)z(m)^*] \simeq 0 , \quad n\neq m. \tag{4.7}$$



$$E[z(0)z(0)^*] = \int_{-\infty}^{\infty} N(x\omega_0)\text{sinc}^2(x-0)dx$$

Variance



$$E[z(0)z(2)^*] = \int_{-\infty}^{\infty} N(x\omega_0)\text{sinc}(x-0)\text{sinc}(x-2)dx$$

Covariance

Figure 4.1. Variance and Covariance of the Fourier Coefficients.

Essentially, increasing the observation interval leads to a smoothing of the power spectral density function when it is written as a function of $x\omega_0$ (as $T \to \infty$, $\omega_0 = 2\pi/T \to 0$). Thus, for a given random process, (4.6) and (4.7) may be considered valid if the observation interval is chosen long enough so that the power spectrum $N(x\omega_0)$ is relatively smooth with respect to increments in ω_0.

A particular power spectrum often assumed is that of bandlimited white Gaussian noise

$$N(\omega) = \begin{cases} No/2 , & |\omega| < 2\pi W \\ 0 , & \text{otherwise} \end{cases} \tag{4.8}$$

where W is the bandwidth in Hertz. It is interesting to observe how the elements of the covariance matrix associated with such a spectrum change as T is allowed to increase. Appendix B contains a set of five such matricies which represent successive doubling of the observation length (Marshall, 1973). The first matrix is for an observation length such that $W = .5\omega_0/2\pi$ or $T = .5/W$; the last matrix is for an observation length such that $W = 8\omega_0/2\pi$ or $T = 8/W$. As T increases, the matricies become progressively more diagonal in form. Roughly, a condition on the observation length can be established on the basis of the last matrix. With reference to the unit height spectrum, for at least 80% of the Fourier coefficients ($0 \leq n \leq N$) to have the following properties

$$\begin{aligned} &(1) \quad .95 \leq E[z(n)z(n)^*] \\ &(2) \quad |E[z(n)z(m)^*]| \leq .02 \end{aligned} \tag{4.9}$$

then $2WT > 16$. Note that this is on the order of the usually assumed condition arising in the uniformly spaced time samples approach to random process representation (i.e., $2WT >> 1$).

Covariance Arising From a Vector Random Process

The previous section dealt only with the covariance between Fourier coefficients representing the time waveform observed at a single element. Presumably, the observation length will be chosen so that (4.6) and (4.7) may be assumed valid. Now, let $\underline{z}(t)$ be a vector of sample functions from the zero mean stationary vector random process observed as the collection of outputs from all the array elements. The Fourier coefficients for these time waveforms will be defined as in Chapter III

$$z_k(n) = \left(\frac{1}{T}\right)^{1/2} \int_{-T/2}^{T/2} z_k(t)\exp(-jn\omega_0 t)\, dt \qquad (4.10)$$

where $\omega_0 = \frac{2\pi}{T}$

and k = the array element index.

Since the relationship between the coefficients representing a single element's output has already been discussed, it remains to consider pairs of Fourier coefficients arising from two different elements.

Clearly, if the random processes observed at all the array elements are independent of one another, their respective collections of Fourier coefficients also will be independent. Such will be the case when we consider independent sensor noise in the problem formulations of Chapter VI. A different situation arises when the noise field contains an additive directional noise component. Now, a portion of the random process observed at one element simply will be a time delayed version of that observed at another. Given that the observation length has been chosen long enough so that (4.6) and (4.7) may be assumed valid, then any two Fourier coefficients at different frequency indicies will be approximately uncorrelated. However, coefficients at the same frequency index will be

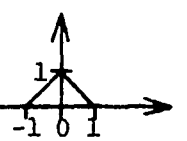
related as shown in Appendix C. Summarizing those results

$$E[z_\ell(n)z_k(n)^*|\tau_{\ell k}] = \exp[-jn\omega_0\tau_{\ell k}]\cdot\int_{-\infty}^{\infty} D(x\omega_0)\mathrm{sinc}^2(x-n)\exp[-j(x-n)(-\frac{\tau_{\ell k}}{T})2\pi]dx \tag{4.11}$$

where $\tau_{\ell k}$ is the time delay between the ℓ^{th} and k^{th} elements of the directional noise component d(t) and $D(x\omega_0)$ is its power spectral density function. Note that if the reception situation is such that plane waves are incident across a uniformly spaced linear array, then $\tau_{\ell k} = (k-\ell)\tau_n$ where τ_n is the time delay of d(t) between adjacent elements. As before, the covariance of the Fourier coefficients can be pictured as the area under a curve. The real part of the integrand in (4.11) is illustrated graphically in Figure 4.2 for the case $\tau_{\ell k}/T = 1/4$. Two conclusions may be drawn

(1) As long as $D(x\omega_0)$ is approximately constant within $2\omega_0$ or $3\omega_0$ either side of $n\omega_0$, then

$$E[z_\ell(n)z_k(n)^*|\tau_{\ell k}] \approx \exp(-jn\omega_0\tau_{\ell k}) \cdot D(n\omega_0) \cdot \{$$



$$\} \tag{4.12}$$

(2) As long as $D(x\omega_0)$ is approximately constant over the interval where the product $\mathrm{sinc}(x-n)\cdot\mathrm{sinc}(x-m)$ has appreciable value, then

$$E[z_\ell(n)z_k(m)^*|\tau_{\ell k}] \approx 0 \quad , \quad n \neq m . \tag{4.13}$$

Thus, for a given additive directional noise component, (4.12) and (4.13) may be considered valid if the observation interval is chosen long enough so that the power spectrum $D(x\omega_0)$ is relatively smooth with respect to increments in ω_0.

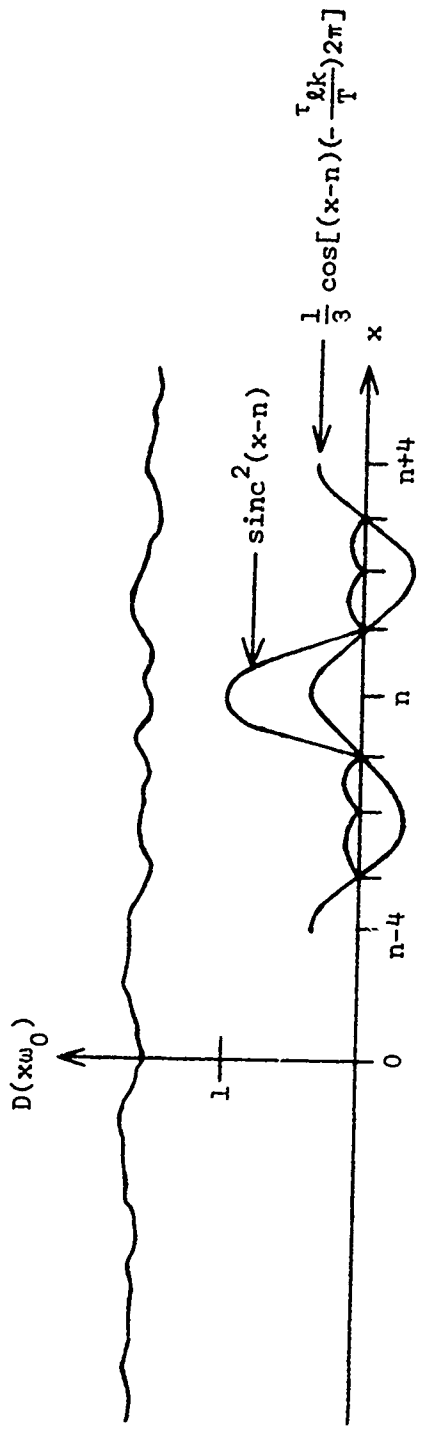


Figure 4.2. Real Part of the Integrand in (4.11).
$\tau_{\ell k}/T = 1/4$

Covariance Between Adjacent Observation Vectors

When the total observation period is broken into several smaller incremental periods each of length T, a sequence of observation vectors will result. The real random processes observed at the output of each array element will be assumed zero mean and stationary. Assuming T is chosen long enough so that (4.6) and (4.7) may be considered valid, then two Fourier coefficients at different frequency indicies and in adjacent observation vectors will be approximately uncorrelated. Coefficients at the same frequency index are related as discussed in Appendix D. Summarizing those results for the case of independent sensor noise

$$E[z_k^i(n)z_k^{i+1}(n)^*] = \int_{-\infty}^{\infty} N(x\omega_0)\text{sinc}^2(x-n)\exp[-j(x-n)2\pi]\,dx \qquad (4.14)$$

where $\underline{z}^i$ and $\underline{z}^{i+1}$ are adjacent observation vectors and $N(x\omega_0)$ is the power spectral density function of the independent noise. And, for an additive directional noise component

$$E[z_\ell^i(n)z_k^{i+1}(n)^*|\tau_{\ell k}]=\exp[-jn\omega_0\tau_{\ell k}]\int_{-\infty}^{\infty} D(x\omega_0)\text{sinc}^2(x-n)\exp[-j(x-n)(1-\frac{\tau_{\ell k}}{T})2\pi]dx \qquad (4.15)$$

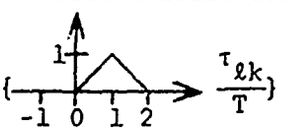
where $D(x\omega_0)$ is the power spectral density function of the directional noise. Two conclusions may be drawn

(1) As long as $N(x\omega_0)$ is approximately constant within $2\omega_0$ or $3\omega_0$ either side of $n\omega_0$, then for independent noise

$$E[z_k^i(n)z_k^{i+1}(n)^*] \approx 0\,. \qquad (4.16)$$

(2) As long as $D(x\omega_0)$ is approximately constant within $2\omega_0$ or $3\omega_0$ either side of $n\omega_0$, then for an additive direction noise component

$$E[z_\ell^i(n)z_k^{i+1}(n)^*|\tau_{\ell k}] \approx \exp(-jn\omega_0\tau_{\ell k}) \cdot D(n\omega_0) \cdot$$



(4.17)

Thus, (4.16) and (4.17) may be considered valid if the incremental observation length T is chosen long enough so that the power spectra $N(x\omega_0)$ and $D(x\omega_0)$ are relatively smooth with respect to increments in ω_0. Under these conditions plus the additional requirement that $\tau_{\ell k} << T$, Fourier coefficients within adjacent observation vectors will have a parameter conditional covariance of approximately zero.

The covariance expressions discussed in this chapter are summarized in Table 4.1. Both the exact expressions and those conditioned on a sufficiently long incremental observation length are given.

Covariance Between the Fourier Coefficients

Let $z_k(t) = n_k(t) + d_0(t - \tau_{0k})$

where $z_k(t)$ = output observed at the k^{th} element

$n_k(t)$ = independent noise component at the k^{th} element

$d_0(t)$ = directional noise component at the 0^{th} elemen.

$\tau_{\ell k}$ = time delay of directional component between the ℓ^{th} and k^{th} elements.

A. Covariance Within a Single Observation Vector $\underline{z}^i$

$$E[z_\ell^i(n)z_k^i(m)^*|\tau_{\ell k}]$$

$$=\exp[-j\omega_0\tau_{\ell k}]\cdot\int_{-\infty}^{\infty}[N(x\omega_0)\cdot\delta_{\ell k}+D(x\omega_0)]\text{sinc}(x-n)\text{sinc}(x-m)\exp[-j(x-m)(-\frac{\tau_{\ell k}}{T})2\pi]\,dx$$

where $$\delta_{\ell k} = \begin{cases} 1 , & \ell=k \\ 0 , & \ell\neq k \end{cases}$$

$N(x\omega_0)$ = power spectral density function of the independent noise component

$D(x\omega_0)$ = power spectral density function of the directional noise component.

B. Covariance Between Adjacent Observation Vectors $\underline{z}^i$ and $\underline{z}^{i+1}$

$$E[z_\ell^i(n)z_k^{i+1}(m)^*|\tau_{\ell k}]$$

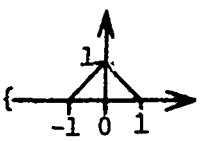
$$=\exp[-jm\omega_0(T-\tau_{\ell k})]\cdot\int_{-\infty}^{\infty}[N(x\omega_0)\cdot\delta_{\ell k}+D(x\omega_0)]\text{sinc}(x-n)\text{sinc}(x-m)$$

$$\exp[-j(x-m)(1 - \frac{\tau_{\ell k}}{T})2\pi]\,dx .$$

Table 4.1

Now, assume that the incremental observation period is chosen long enough so that the power spectra $N(x\omega_0)$ and $D(x\omega_0)$ are relatively smooth with respect to increments in ω_0.

C. Covariance Within a Single Observation Vector $\underline{z}^i$

$$E[z_\ell^i(n)z_k^i(m)^*|\tau_{\ell k}] \simeq \exp[-jm\omega_0\tau_{\ell k}]\cdot[N(n\omega_0)\cdot\delta_{\ell k}+D(n\omega_0)]\cdot \{$$  $$\frac{\tau_{\ell k}}{T}\}\cdot\delta_{nm}.$$

D. Covariance Between Adjacent Observation Vectors $\underline{z}^i$ and $\underline{z}^{i+1}$

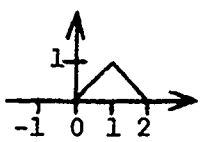
$$E[z_\ell^i(n)z_k^{i+1}(m)^*|\tau_{\ell k}] \simeq [-jm\omega_0(T-\tau_{\ell k})]\cdot[0+D(n\omega_0)]\cdot \{$$  $$\frac{\tau_{\ell k}}{T}\}\cdot\delta_{nm}.$$

Table 4.1 (continued)

Chapter V

TIME SEQUENTIAL ARRAY PROCESSORS

The general design equations for sequential data processing were given in (2.8), (2.9) and (2.10). No specific restrictions were made on the nature of the observation vector (i.e., $\underline{R}_K$ could be a collection of uniformly spaced time samples of an ongoing waveform or perhaps the Fourier coefficients of a fixed length signal). When the R_i are uniformly spaced time samples, sequential processing leads to a natural learning feature with time as more observations are processed. The frequency domain analog for a fixed length signal exhibits learning, not in time, but in frequency sequence. The thrust of this chapter will be to propose a model of waveform representation which breaks the total observation length into a sequence of fixed length incremental observation intervals and then to analyze the consequences in terms of time sequential processing utilizing Fourier coefficients.

Time Sequential Structure*

Consider the vector of stationary random processes $\underline{z}(t)$ as in (3.12) which is the vector of observed outputs from each array element. Based on some criterion (as yet unchosen), pick an incremental observation period and break $\underline{z}(t)$ into time sequences of length T_{inc}. The corresponding vector of Fourier coefficients representing the i^{th} period will be $\underline{z}^i$ as in (3.15). Thus, for a total observation length of $T = L \cdot T_{inc}$, $\underline{z}(t)$ will be represented by L vectors of Fourier coefficients, $(\underline{z}^1, \ldots, \underline{z}^L)$.

In the formation of the likelihood ratio, the marginal distribution of the observables conditional to each hypothesis is needed. Our observables are now the L vectors $(\underline{z}^1,\ldots,\underline{z}^L)$. Suspending the conditioning to H_1 and H_0

$$p(\underline{z}^1,\ldots,\underline{z}^L) = \prod_{i=1}^{L} p(\underline{z}^i|\underline{z}^{i-1},\ldots,\underline{z}^1)$$

$$= \int_\theta \prod_{i=1}^{L} p(\underline{z}^i|\underline{z}^{i-1},\ldots,\underline{z}^1,\underline{\theta})p(\underline{\theta})d\underline{\theta} . \qquad (5.1)$$

Assuming parameter conditional independence of the $\underline{z}^i$

$$p(\underline{z}^i|\underline{z}^{i-1},\ldots,\underline{z}^1,\underline{\theta}) = p(\underline{z}^i|\underline{\theta}) . \qquad (5.2)$$

Substituting (5.2) into (5.1)

$$p(\underline{z}^1,\ldots,\underline{z}^L) = \int_\theta \prod_{i=1}^{L} p(\underline{z}^i|\underline{\theta})p(\underline{\theta})d\underline{\theta} . \qquad (5.3)$$

Lastly, applying Bayes' rule L times to the integrand in (5.3)

$$p(\underline{z}^1,\ldots,\underline{z}^L) = \prod_{i=1}^{L} \int_\theta p(\underline{z}^i|\underline{\theta})p(\underline{\theta}|\underline{z}^{i-1},\ldots,\underline{z}^1)d\underline{\theta} \qquad (5.4)$$

where $p(\underline{\theta}|\underline{z}^{i-1},\ldots,\underline{z}^1)$ is the updated version of the *a priori* probability density function of $\underline{\theta}$

$$p(\underline{\theta}|\underline{z}^{i-1},\ldots,\underline{z}^1) = \frac{p(\underline{z}^{i-1}|\underline{\theta})p(\underline{\theta}|\underline{z}^{i-2},\ldots,\underline{z}^1)}{p(\underline{z}^{i-1}|\underline{z}^{i-2},\ldots,\underline{z}^1)} . \qquad (5.5)$$

Sequential processing in time as summarized in (5.4) and (5.5) now can be accomplished using observables taken in the frequency domain to form the marginal distributions required to calculate the likelihood ratio

$$\Lambda(\underline{z}^1,\ldots,\underline{z}^L) = \frac{p(\underline{z}^1,\ldots,\underline{z}^L|H_1)}{p(\underline{z}^1,\ldots,\underline{z}^L|H_0)} . \qquad (5.6)$$

Figure 5.1 illustrates the corresponding array processor structure. Note that at the end of each incremental observation period, the prior knowledge for that period is updated under H_1 and H_0 to reflect the processing of an additional increment of data. These updated densities then are used as prior knowledge for the next incremental observation period.

The essential advantages of structuring frequency domain array processors in a time sequential fashion are

(1) No need to specify a priori the total observation length. Normally, the number and location of frequency samples is dependent on the total observation length through $\omega_0 = 2\pi/T$. Operating sequentially, we fix the incremental observation period T_{inc} and thereby fix $\omega_0 = 2\pi/T_{inc}$ for any $T = L \cdot T_{inc}$, L an integer. Thus, the processor structure remains fixed for any total observation length.

(2) Elimination of a linearly growing memory requirement. As mentioned in (1), the number and location of frequency samples is dependent on T for the one shot processor. When T doubles, the number of frequency samples collected for a given band-limited observed random process doubles, etc. Thus, twice as many Fourier coefficients must be remembered and manipulated to calculate the likelihood ratio. The sequential structure eliminates this difficulty by permitting the processing of (1/L) the total number of Fourier coefficients at the end of each incremental observation period.

(3) The learning or adaptive features in time which take place naturally with a time sequential Bayesian processor.

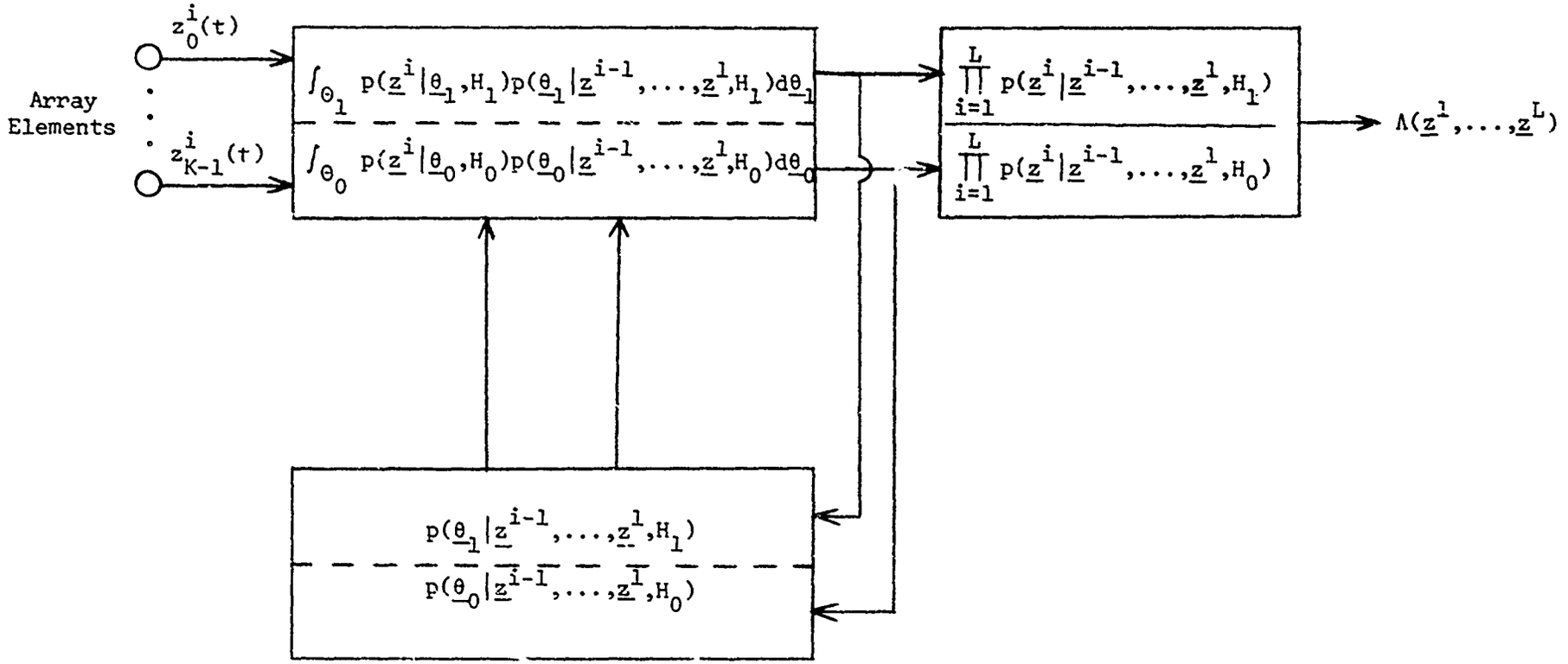


Figure 5.1. Time Sequential Array Processor.

Single Frequency/Conjugate Prior Decomposition

Now, make the additional assumption of parameter conditional independence between Fourier coefficients at different frequency indicies and within the same observation vector

$$p(\underline{z}^i(0),\ldots,\underline{z}^i(N)|\underline{\theta}) = p(\underline{z}^i(0)|\underline{\theta})\ldots p(\underline{z}^i(N)|\underline{\theta}) \; . \tag{5.7}$$

The resulting separation of the joint probability density function of the observables in frequency modifies (5.3) to become

$$\int_\Theta \underbrace{\boxed{p(\underline{z}^1(0)|\underline{\theta})\ldots p(\underline{z}^1(N)|\underline{\theta})}}_{p(\underline{z}^1|\underline{\theta})} \ldots \underbrace{\boxed{p(\underline{z}^L(0)|\underline{\theta})\ldots p(\underline{z}^L(N)|\underline{\theta})}}_{p(\underline{z}^L|\underline{\theta})} p(\underline{\theta})\, d\underline{\theta}$$

1^{st} observation period $\qquad$ L^{th} observation period

$$= \int_\Theta \underbrace{\boxed{p(\underline{z}^1(0)|\underline{\theta})\ldots p(\underline{z}^L(0)|\underline{\theta})}}_{0^{th}\text{ frequency}} \ldots \underbrace{\boxed{p(\underline{z}^1(N)|\underline{\theta})\ldots p(\underline{z}^L(N)|\underline{\theta})}}_{N^{th}\text{ frequency}} p(\underline{\theta})\, d\underline{\theta} \tag{5.8}$$

where $p(\underline{z}^i(n)|\underline{\theta})$ is the conditional density of the n^{th} frequency component of the i^{th} observation vector as in (3.14).

Assume that all the conditional densities in (5.8) admit sufficient statistics of finite dimension for $\underline{\theta}$ (see Theorem 2.1, Chapter II). Let

$$p(\underline{\theta}) = r_0(\underline{\theta})p_0(\underline{\theta}) \tag{5.9}$$

where $p_0(\underline{\theta})$ is a natural conjugate prior under $p(\underline{z}^i(0)|\underline{\theta})$ and $r_0(\underline{\theta})$ is as in (2.16). Then, utilizing Bayes' rule

$$\boxed{p(\underline{z}^1(0)|\underline{\theta})\ldots p(\underline{z}^L(0)|\underline{\theta})}\; p(\underline{\theta}) = \boxed{p(z^1(0)|\underline{\theta})\ldots p(z^L(0)|\underline{\theta})}\; r_0(\underline{\theta})\; p_0(\underline{\theta})$$

$$= p(\underline{z}^1(0),\ldots,\underline{z}^L(0);p_0(\underline{\theta}))p(\underline{\theta}|\underline{z}^1(0),\ldots,\underline{z}^L(0);p_0(\underline{\theta}))\; r_0(\underline{\theta}) \tag{5.10}$$

where $p(\underline{z}^1(0),\ldots,\underline{z}^L(0);p_0(\underline{\theta}))$ is the marginal distribution of $(\underline{z}^1(0),\ldots,\underline{z}^L(0))$ and $p(\underline{\theta}|\underline{z}^L(0),\ldots,\underline{z}^L(0);p_0(\underline{\theta}))$ is the *a posteriori* distribution of $\underline{\theta}$ based on a prior of $p_0(\underline{\theta})$. Next, let

$$p(\underline{\theta}|\underline{z}^1(0),\ldots,\underline{z}^L(0);p_0(\underline{\theta})) = r_1(\underline{\theta})p_1(\underline{\theta}) \tag{5.11}$$

where $p_1(\underline{\theta})$ is a natural conjugate prior under $p(\underline{z}^i(1)|\underline{\theta})$ and $r_1(\underline{\theta})$ is as in (2.16). Continuing the alternate application of Bayes' rule and the incorporation of natural conjugate priors for each of the N+1 frequencies, (5.8) becomes

$$\begin{aligned} p(\underline{z}^1,\ldots,\underline{z}^L) &= p(\underline{z}^1(0),\ldots,\underline{z}^L(0);p_0(\underline{\theta}))\ldots p(\underline{z}^1(N),\ldots,\underline{z}^L(N);p_N(\underline{\theta})) \\ &\quad \cdot \int_\Theta r_0(\underline{\theta})\ldots r_N(\underline{\theta})p(\underline{\theta}|\underline{z}^1(N),\ldots,\underline{z}^L(N);p_N(\underline{\theta}))d\underline{\theta} \\ &= \{\prod_{n=0}^{N} p(\underline{z}^1(n),\ldots,\underline{z}^L(n);p_n(\theta))\} \\ &\quad \cdot \int_\Theta \{\prod_{n=0}^{N} r_n(\underline{\theta})\}p(\underline{\theta}|\underline{z}^1(N),\ldots,\underline{z}^L(N);p_N(\underline{\theta}))d\underline{\theta}\,. \end{aligned} \tag{5.12}$$

The notation

$$p(\underline{\theta}|\underline{z}^1(n),\ldots,\underline{z}^L(n);p_n(\underline{\theta})) = r_{n+1}(\underline{\theta})p_{n+1}(\underline{\theta}) \tag{5.13}$$

has been used to indicate the relationship between the *a posteriori* pdf of $\underline{\theta}$ after $(\underline{z}^1(n),\ldots,\underline{z}^L(n))$ has been processed and the *a priori* pdf of $\underline{\theta}$ used for processing $(\underline{z}^1(n+1),\ldots,\underline{z}^L(n+1))$. Essentially, the data dependency has been concentrated in $r_{n+1}(\underline{\theta})$ thus allowing freedom in the choice of $p_{n+1}(\underline{\theta})$. Figure 5.2 illustrates the general array processor structure implied by (5.12). The single frequency sequential/conjugate blocks are the realizations of $p(\underline{z}^1(n),\ldots,\underline{z}^L(n),p_n(\underline{\theta}))$; n = 0,...,N. Each is an independent time sequential processor similar to Figure 5.1 for a single frequency index using its own natural conjugate prior. Only

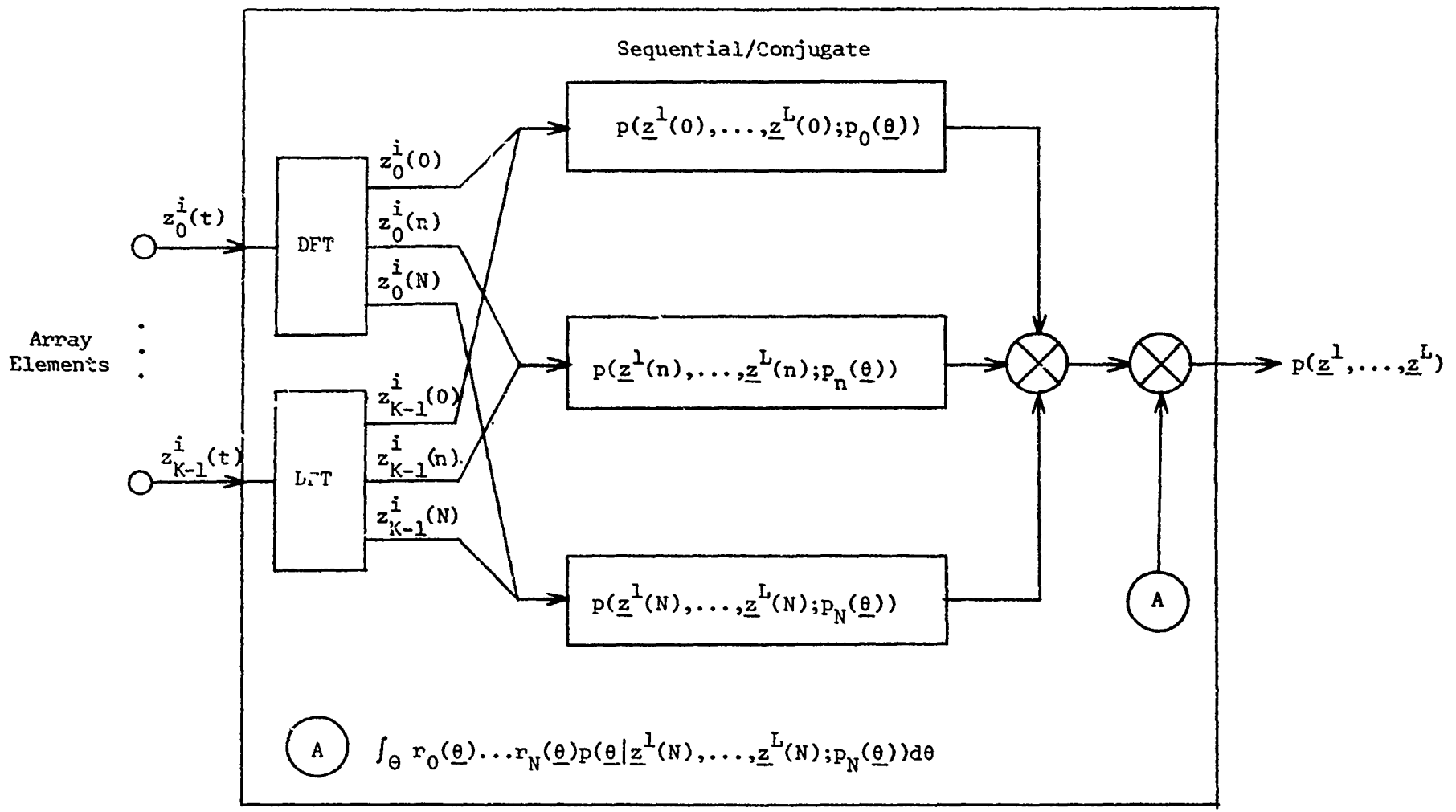


Figure 5.2. General Time Sequential Array Processor.

$$\Lambda(\underline{z}^1,\ldots,\underline{z}^L) = \frac{p(\underline{z}^1,\ldots,\underline{z}^L|H_1)}{p(\underline{z}^1,\ldots,\underline{z}^L|H_0)}$$

the n^{th} Fourier coefficient from each array element is utilized by the n^{th} block. The integral provides the correct matching between the outputs of the N+1 blocks and the actual a priori knowledge contained in $r_0(\underline{\theta})$ to yield the desired marginal density of all the observables. Note the particular form of this integral when uniform priors are admissible (and so chosen) for the natural conjugate priors $p_0(\underline{\theta}),\ldots,p_N(\underline{\theta})$

$$\int_\Theta \{\prod_{n=0}^{N} r_n(\underline{\theta})\}\, p(\underline{\theta}|\underline{z}^1(N),\ldots,\underline{z}^L(N);p_N(\underline{\theta}))d\underline{\theta}$$

$$= C \int_\Theta p(\underline{\theta})\{\prod_{n=0}^{N} p(\underline{\theta}|\underline{z}^1(n),\ldots,\underline{z}^L(n);p_n(\underline{\theta})\}d\underline{\theta} \qquad (5.14)$$

where C is a constant. The integral becomes simply a weighting of all the a posteriori single frequency/conjugate densities by the true a priori knowledge. To form the likelihood ratio, the marginal density of the observables conditional to H_1 is divided by the marginal density of the observables conditional to H_0 as in (5.6).

The decomposition s ggested by (5.12) is advantageous for the following reasons

(1) Conceptually, the general array processor is seen to separate into a few well defined components. In addition, since all the sequential/conjugate blocks are mathematically similar, only the n^{th} frequency block must be derived in detail.

(2) Practically, a majority of the general array processor can be designed without knowing the actual a priori knowledge. Furthermore, the inherent mathematical tractability of natural conjugate priors may be beneficial in the development of the single frequency sequential/conjugate blocks.

Criterion for the Selection of T_{inc}^{*}

Derivation of the sequential optimal array processors in this chapter relied on the two assumptions of parameter conditional independence made in (5.2) and (5.7). Independence in time led to (5.3), and the addition of independence in frequency permitted the decomposition shown in (5.8). For jointly Gaussian random variables, zero covariance implies statistical independence. Thus (assuming the observables have a parameter conditional expected value of zero) (5.2) is valid when

$$E[z_{\ell}^{i}(n)z_{k}^{i+1}(m)^{*}|\underline{\theta}] = 0 \tag{5.15}$$

and (5.7) is valid when

$$E[z_{\ell}^{i}(n)z_{k}^{i}(m)^{*}|\underline{\theta}] = 0 \ , \ n \neq m \ . \tag{5.16}$$

In this section, sufficient conditions for the selection of an incremental observation period T_{inc} are stated which insure the approximate validity of (5.2) and (5.7) for a particular class of problems. In this class, the noise field is Gaussian and consists of a component independent from sensor to sensor plus an additive directional component. The discussion in Chapter IV relates the covariance properties of the Fourier coefficients to observation interval length for such a noise field. The covariance expression corresponding to (5.15) can be found in Table 4.1 B and (5.16) in Table 4.1 A.

Sufficient conditions for the selection of a T_{inc} are as follows

(1) The incremental observation period shall be chosen long enough so that the noise power spectral density function when written as a function of $x\omega_0$ is relatively smooth with respect to increments in $\omega_0 = 2\pi/T_{inc}$.

(2) The incremental observation period shall be chosen long enough so that the transit time across the array for the directional noise wavefront τ_{OK} is much less than T_{inc}.

The condition in (1) insures that any two Fourier coefficients at different frequency indicies, be they within the same observation vector $\underline{z}^i$ or within adjacent observation vectors $\underline{z}^i$ and $\underline{z}^{i+1}$, will have a parameter conditional covariance of approximately zero

$$E[z_\ell^i(n)z_k^i(m)^*|\underline{\theta}] \simeq 0 \quad , \quad n \neq m \qquad \text{(Table 4.1 C)}$$

$$E[z_\ell^i(n)z_k^{i+1}(m)^*|\underline{\theta}] \simeq 0 \quad , \quad n \neq m \, . \qquad \text{(Table 4.1 D)}$$

The additional condition in (2) insures that two Fourier coefficients at the same frequency index and within adjacent observation vectors $\underline{z}^i$ and $\underline{z}^{i+1}$ also will have a parameter conditional covariance of approximately zero

$$E[z_\ell^i(n)z_k^{i+1}(n)^*|\underline{\theta}] \simeq 0 \quad , \quad \tau_{\ell k} \leq \tau_{OK} << T_{inc} \, .$$

(Table 4.1 D)

Note that both conditions (1) and (2) are necessary to approximate (5.15) while only condition (1) is required to approximate (5.16). Once a T_{inc} is chosen which permits the assumption of (5.2), the parameter conditional independence in (5.7) follows automatically when the noise field is within the particular class being considered here.

---

*A specialization of notation occurs in this Chapter. T_{inc} is used to denote explicitly the incremental observation period length while T is reserved to denote the total observation period length.

Chapter VI

THREE SPECIFIC PROBLEMS AND THEIR CONDITIONAL LIKELIHOOD RATIOS

The focus of this chapter will be on the derivation of the likelihood ratio processors for three specific problems. In each, directional uncertainty arises in either the location of a signal or noise source. Our observables will consist of the collection of Fourier coefficients $\underline{z}$ representing the bandlimited time waveforms observed at all of the array elements. The primary goal here is to derive the single frequency parameter conditional joint density expressions under H_1 and H_0. These are the fundamental components required by all four of the optimal array processor structures discussed thus far. Also considered are the essential features of the sequential/conjugate blocks indicated in Figure 5.2 which evolve out of the mathematics of each problem. Lastly, the incorporation of a conjugate prior into each block and the integral which results is discussed.

Uncertain Direction: General Formulation

The problems considered in this chapter involve either signal or noise location uncertainty. As mentioned in Chapter II, likelihood ratio expressions for such problems were originally pursued by Adams (Adams, 1973). An array of K uniformly spaced elements on a straight line is assumed with the zeroth element being the right-most sensor. The processor will be asked to decide between the two mutually exclusive and exhaustive hypotheses H_1 that the time waveforms observed at the elements consist of signal plus noise and H_0 that they consist of noise alone.

When noise alone is observed, the array element outputs are sample functions of zero mean, stationary Gaussian random processes. One component of the noise process is assumed to be spatially uncorrelated (i.e., independent from sensor to sensor). The power spectral density function of this component $N(\omega)$ is assumed to be the same at each element. The second component of the noise process is due to a directional Gaussian noise source in the far field of the array. The power spectral density function of this spatially correlated component is denoted $D(\omega)$. Both components are assumed bandlimited. Utilizing the notation of Chapter III and choosing an observation period sufficiently long for the expression in Table 4.1 C to be assumed valid

$$p(\underline{z}|H_0) = \int_{\omega_0\tau_n} \prod_{n=0}^{N} \{\frac{1}{\pi^K |\underline{Q}_0(n,\tau_n)|} \exp[-\underline{z}(n)^* \underline{Q}_0(n,\tau_n)^{-1} \underline{z}(n)]\} p(\omega_0\tau_n) d\omega_0\tau_n \tag{6.1}$$

where

$$\underline{Q}_0(n,\tau_n)^{-1} = N^{-1}\underline{I} - \frac{D\,\underline{v}\,\underline{v}^*}{N(N+K\cdot D)} \tag{6.2}$$

$$N = N(n\omega_0) \text{ and } D = D(n\omega_0)$$

$$\underline{v}(n,\tau_n)^* = [1, \exp(jn\omega_0\tau_n), \ldots, \exp(j(K-1)n\omega_0\tau_n)] \tag{6.3}$$

and

$$\omega_0 = \frac{2\pi}{T}\,, \quad T \text{ the observation period length.}$$

Uncertainty in location is reflected in the parameter τ_n which is the time delay of the directional noise between adjacent elements. Our a priori knowledge on this parameter is summarized by the probability density function $p(\omega_0\tau_n)$. Since the covariance matrix is of the special form

$$\underline{Q}_0(n,\tau_n) = N\,\underline{I} + D\,\underline{v}\,\underline{v}^* \tag{6.4}$$

where $\underline{I}$ is the K-dimensional identity matrix and $\underline{v}(n,\tau_n)$ is the "pointing vector" in the direction of the noise, its inverse in (6.2) is given by Bartlett (Bartlett, 1951). A detailed discussion of this result is given in Appendix E.

The signal consists of a Gaussian random process with a known time varying mean. When signal plus noise is observed, the array element outputs consist of time delayed versions of this signal process added to the noise processes described above. The random component of the bandlimited signal process has a power spectral density function denoted by $S(\omega)$ and the deterministic component is represented by the Fourier coefficients $b_0(n)$ at the zeroth array element. Again, utilizing the notation of Chapter III and choosing an observation period sufficiently long for the expression in Table 4.1 C to be assumed valid

$$p(\underline{z}|H_1) = \int_{\omega_0\tau_s}\int_{\omega_0\tau_n}\prod_{n=0}^{N}\{\frac{1}{\pi^K|\underline{Q}_1(n,\tau_s,\tau_n)|}$$

$$\cdot \exp[-(\underline{z}(n)-b_0(n)\underline{u}(n,\tau_s))^*\underline{Q}_1(n,\tau_s,\tau_n)^{-1}(\underline{z}(n)-b_0(n)\underline{u}(n,\tau_s))]\}$$

$$\cdot p(\omega_0\tau_s,\omega_0\tau_n)d\omega_0\tau_n d\omega_0\tau_s \qquad (6.5)$$

where $\underline{Q}_1(n,\tau_s,\tau_n)^{-1} = N^{-1}I - \frac{D(N+K\cdot S)\underline{v}\underline{v}^*+S(N+K\cdot D)\underline{u}\underline{u}^*-2\cdot D\cdot S\ \mathrm{Re}\{\underline{u}\underline{u}^*\underline{v}\underline{v}^*\}}{N[(N+K\cdot D)(N+K\cdot S)-D^2S\ \underline{u}^*\underline{v}\underline{v}^*\underline{u}]}$ (6.6)

$N = N(n\omega_0)$, $D = D(n\omega_0)$, and $S = S(n\omega_0)$

$$\underline{v}(n,\tau_n)^* = [1,\ \exp(jn\omega_0\tau_n),\ldots,\exp(j(K-1)n\omega_0\tau_n)]$$

$$\underline{u}(n,\tau_s)^* = [1,\ \exp(jn\omega_0\tau_s),\ldots,\exp(j(K-1)n\omega_0\tau_s)] \qquad (6.7)$$

and $\omega_0 = \frac{2\pi}{T}$, T the observation period length.

Uncertainty in the signal location is reflected in the parameter τ_s which is the time delay of the signal between adjacent elements. Under this hypothesis, the covariance matrix is of the form

$$\underline{Q}_1(n,\tau_s,\tau_n) = N\,\underline{I} + D\,\underline{v}\underline{v}^* + S\,\underline{u}\underline{u}^* \tag{6.8}$$

whose inverse in (6.6) also is discussed in Appendix E.

The likelihood ratio is by definition the ratio of (6.5) and (6.1)

$$\Lambda(\underline{z}) \triangleq \frac{p(\underline{z}|H_1)}{p(\underline{z}|H_0)}. \tag{6.9}$$

Three Problems

The preceding section provided the general formulation of the array processor when there is location uncertainty. Note how the integrands in (6.5) and (6.1) separate into their single frequency constituents. These single frequency conditional densities will be derived for three specific problems

(1) Gaussian Signal of Uncertain Direction in Gaussian Noise (GUD),

(2) Signal Known Except for Direction in Noise with an Additive Directional Noise Component of Known Direction (SKED in NKD),

(3) Signal Known Exactly in Noise with an Additive Directional Noise Component of Uncertain Direction (SKE in NUD).

GUD

Under H_0

$$p(\underline{z}(n)|H_0) = \frac{1}{\pi^K|\underline{Q}_0(n)|}\exp[-\underline{z}(n)^*\underline{Q}_0(n)^{-1}\underline{z}(n)] \tag{6.10}$$

where

$$\underline{Q}_0(n) = N\,\underline{I} \tag{6.11}$$

$$N = N(n\omega_0)$$

$$|\underline{Q}_0(n)| = N^K \tag{6.12}$$

and

$$\underline{Q}_0(n)^{-1} = N^{-1}\underline{I} \,. \tag{6.13}$$

Under H_1

$$p(\underline{z}(n)|\omega_0\tau_s,H_1) = \frac{1}{\pi^K|\underline{Q}_1(n,\tau_s)|} \exp[-\underline{z}(n)^*\underline{Q}_1(n)^{-1}\underline{z}(n)] \tag{6.14}$$

where

$$\underline{Q}_1(n,\tau_s) = N\,\underline{I} + S\,\underline{u}\,\underline{u}^* \tag{6.15}$$

$$N = N(n\omega_0) \text{ and } S = S(n\omega_0)$$

$$\underline{u}(n,\tau_s)^* = [1,.., \exp(jn\omega_0\tau_s),..., \exp(j(K-1)n\omega_0\tau_s)]$$

$$|\underline{Q}_1(n,\tau_s)| = N^{K-1}(K\cdot S+N) \tag{6.16}$$

and

$$\underline{Q}_1(n,\tau_s)^{-1} = N^{-1}\underline{I} - \frac{S\,\underline{u}\,\underline{u}^*}{N(N+K\cdot S)} \,. \tag{6.17}$$

Since the uncertain parameter exists only under H_1, the single frequency conditional likelihood ratio is of interest

$$\Lambda(\underline{z}(n)|\omega_0\tau_s) \triangleq \frac{p(\underline{z}(n)|\omega_0\tau_s,H_1)}{p(\underline{z}(n)|H_0)}$$

$$= \frac{|\underline{Q}_0(n)|}{|\underline{Q}_1(n)|} \exp[-\underline{z}(n)^*(\underline{Q}_1(n,\tau_s)^{-1}-\underline{Q}_0(n)^{-1})\underline{z}(n)]. \tag{6.18}$$

Making the appropriate substitutions from (6.10)-(6.17) into (6.18)

$$\Lambda(\underline{z}(n)|\omega_0\tau_s) = \frac{N}{K\cdot D+N} \exp[\frac{S}{N(N+K\cdot S)} \sum_{i=0}^{K-1} z_i(n)^* z_i(n)]$$

$$\cdot \exp[\frac{S}{N(N+K\cdot S)}\, 2\, \{\sum_{\ell=1}^{K-1} A_\ell \cos(\ell n\omega_0\tau_s+B_\ell)\}] \tag{6.19}$$

where
$$A_\ell \cos B_\ell = \mathrm{Re}[C_\ell]$$
$$A_\ell \sin B_\ell = \mathrm{Im}[C_\ell] \tag{6.20}$$

and
$$C_\ell \triangleq \sum_{i=0}^{K-1-\ell} z_i(n)^* z_{i+\ell}(n) \; . \tag{6.21}$$

The essential features of the processor structure for this problem are illustrated in Figure 6.1.

SKED in NKD

Under H_0

$$p(\underline{z}(n)|H_0) = \frac{1}{\pi^K |\underline{Q}(n)|} \exp[-\underline{z}(n)^* \underline{Q}(n)^{-1} \underline{z}(n)] \tag{6.22}$$

where
$$\underline{Q}(n) = N \underline{I} + D \underline{v}\underline{v}^* \tag{6.23}$$

$$N = N(n\omega_0) \text{ and } D = D(n\omega_0)$$

$$\underline{v}(n)^* = [1, \exp(jn\omega_0\tau_n), \ldots, \exp(j(K-1)n\omega_0\tau_n)]$$

and
$$\underline{Q}(n)^{-1} = N^{-1}\underline{I} - \frac{D \underline{v}\underline{v}^*}{N(N+K\cdot D)} \; . \tag{6.24}$$

Under H_1

$$p(\underline{z}(n)|\omega_0\tau_s, H_1) = \frac{1}{\pi^K |\underline{Q}(n)|} \exp[-(\underline{z}(n)-b_0(n)\underline{u}(n,\tau_s))^* \underline{Q}(n)^{-1}(\underline{z}(n)-b_0(n)\underline{u}(n,\tau_s))] \tag{6.25}$$

where $\underline{u}(n,\tau_s)^* = [1, \exp(jn\omega_0\tau_s), \ldots, \exp(j(K-1)n\omega_0\tau_s)]$.

Since the uncertain parameter exists only under H_1, the single frequency conditional likelihood ratio is of interest

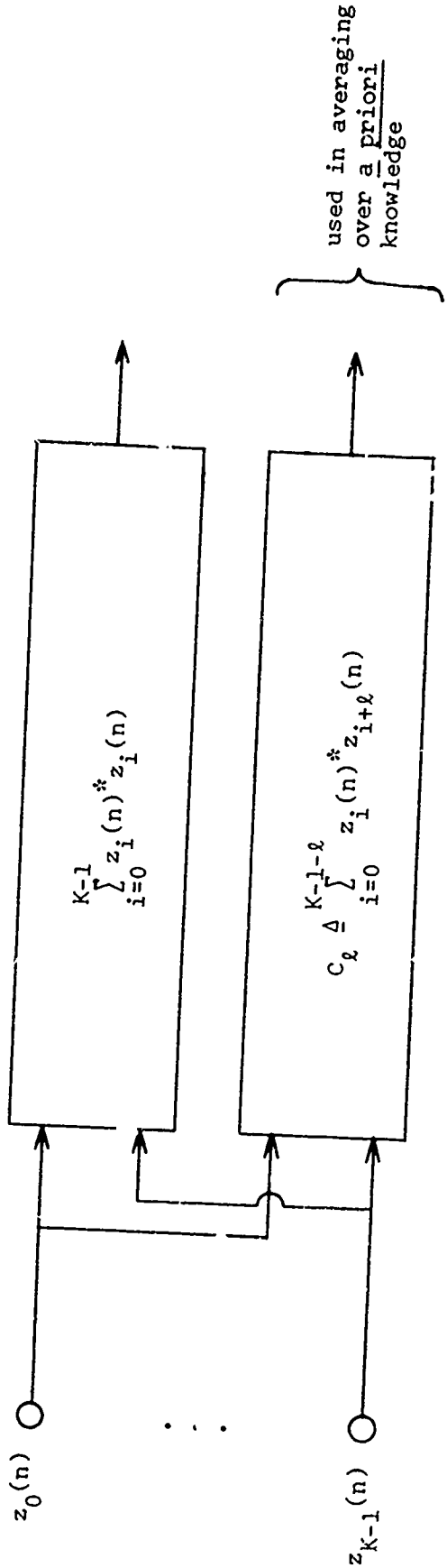


Figure 6.1. Essential Features of the Likelihood Ratio for GUD.

$$\Lambda(\underline{z}(n)|\omega_0\tau_s) \triangleq \frac{p(\underline{z}(n)|\omega_0\tau_s,H_1)}{p(\underline{z}(n)|H_0)}$$

$$= \exp[2\,\mathrm{Re}\{\underline{z}(n)^*\underline{Q}(n)^{-1}b_0(n)\underline{u}(n,\tau_s)\} - |b_0(n)|^2\underline{u}(n,\tau_s)^*\underline{Q}(n)^{-1}\underline{u}(n,\tau_s)]. \tag{6.26}$$

Making the appropriate substitutions from (6.22)-(6.25) into (6.26)

$$\Lambda(\underline{z}(n)|\omega_0\tau_s) = \exp[-\frac{K[N+(K-1)D]}{N(N+K\cdot D)}|b_0(n)|^2$$

$$+ 2\,\mathrm{Re}\{\frac{1}{N}z_0(n)^*b_0(n) - \frac{D}{N(N+K\cdot D)}\sum_{i=0}^{K-1}z_i(n)^*b_0(n)\exp(-jin\omega_0\tau_n)\}]$$

$$\cdot \exp[2\sum_{\ell=1}^{K-1}A_\ell\cos(\ell n\omega_0\tau_s + B_\ell)] \tag{6.27}$$

where

$$A_\ell\cos B_\ell = \mathrm{Re}[G_\ell] - \mathrm{Re}[H_\ell] + \mathrm{Re}[I_\ell]$$

$$A_\ell\sin B_\ell = -\mathrm{Im}[G_\ell] + \mathrm{Im}[H_\ell] - \mathrm{Im}[I_\ell] \tag{6.28}$$

and

$$G_\ell \triangleq \frac{1}{N}z_\ell(n)^*b_0(n) \tag{6.29}$$

$$H_\ell \triangleq \frac{D}{N(N+K\cdot D)}\sum_{i=0}^{K-1}z_i(n)^*b_0(n)\exp(-j(i-\ell)n\omega_0\tau_n) \tag{6.30}$$

$$I_\ell \triangleq \frac{D}{N(N+K\cdot D)}|b_0(n)|^2(K-\ell)\exp(j\ell n\omega_0\tau_n)\,. \tag{6.31}$$

The essential features of the processor structure for this problem are illustrated in Figure 6.2.

SKE in NUD

Under H_0

$$p(\underline{z}(n)|\omega_0\tau_n,H_0) = \frac{1}{\pi^K|\underline{Q}(n,\tau_n)|}\exp[-\underline{z}(n)^*\underline{Q}(n,\tau_n)^{-1}\underline{z}(n)] \tag{6.32}$$

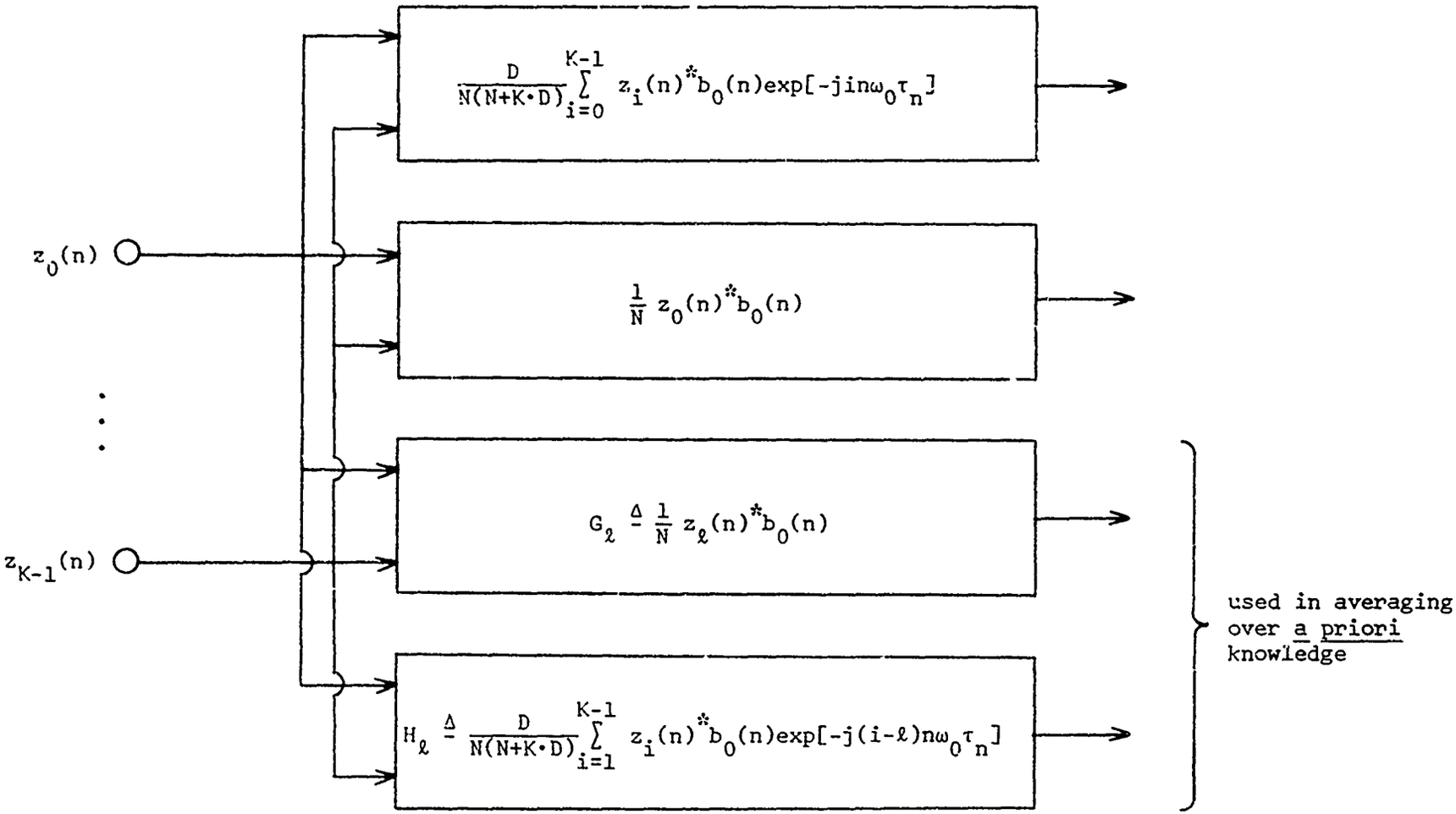


Figure 6.2. Essential Features of the Likelihood Ratio for SKED in NKD.

where $$\underline{Q}(n,\tau_n) = N\,\underline{I} + D\,\underline{v}\,\underline{v}^* \tag{6.33}$$

$$N = N(n\omega_0) \text{ and } D = D(n\omega_0)$$

$$\underline{v}(n,\tau_n)^* = [1,\ \exp(jn\omega_0\tau_n),\ \ldots,\ \exp(j(K-1)n\omega_0\tau_n)]$$

$$|\underline{Q}(n,\tau_n)| = N^{K-1}(K\cdot D+N)$$

and $$\underline{Q}(n,\tau_n)^{-1} = N^{-1}\underline{I} - \frac{D\,\underline{v}\,\underline{v}^*}{N(N+K\cdot D)}\,. \tag{6.34}$$

Under H_1

$$p(\underline{z}(n)|\omega_0\tau_n,H_1) = \frac{1}{\pi^K\,|\underline{Q}(n,\tau_n)|}\exp[-(\underline{z}(n)-b_0(n)\underline{u}(n))^*\underline{Q}(n,\tau_n)^{-1}(\underline{z}(n)-b_0(n)\underline{u}(n))]$$

$$= \frac{1}{\pi^K|\underline{Q}(n,\tau_n)|}\exp[-\underline{z}(n)^*\underline{Q}(n,\tau_n)^{-1}\underline{z}(n)+2\mathrm{Re}\{\underline{z}(n)^*\underline{Q}(n,\tau_n)^{-1}b_0(n)\underline{u}(n)\}$$

$$- |b_0(n)|^2\underline{u}(n)^*\underline{Q}(n,\tau_n)^{-1}\underline{u}(n)] \tag{6.35}$$

where $$\underline{u}(n)^* = [1,\ \exp(jn\omega_0\tau_s),\ \ldots,\ \exp(j(K-1)n\omega_0\tau_s)]\,.$$

Since the uncertain parameter exists under both hypotheses, the conditional joint densities must be retained individually. Making the appropriate substitutions from (6.32)-(6.35) into (6.32) and (6.35)

$$p(\underline{z}(n)|\omega_0\tau_n,H_0) =$$

$$\frac{1}{\pi^K N^{K-1}(K\cdot D+N)}\exp\left[-\frac{[N+(K-1)D]}{N(N+K\cdot D)}\sum_{i=0}^{K-1}z_i(n)^*z_j(n)\right]\cdot\exp\left[\frac{D}{N(N+K\cdot D)}2\sum_{\ell=1}^{K-1}A_\ell^0\cos(\ell n\omega_0\tau_n+B_\ell^0)\right] \tag{6.36}$$

and

$$p(\underline{z}(n)|\omega_0\tau_n,H_1) = \frac{1}{\pi^{KN}K^{-1}(K\cdot D+N)} \exp[-\frac{[N+(K-1)D]}{N(N+K\cdot D)}\sum_{i=0}^{K-1} z_i(n)^* z_i(n)$$

$$+ 2\,\mathrm{Re}\{\frac{N+(K-1)D}{N(N+K\cdot D)}\sum_{i=0}^{K-1} z_i(n)^* b_0(n)\exp(-jin\omega_0\tau_s)\} - \frac{K[N+(K-1)D]}{N(N+K\cdot D)}\,|b_0(n)|^2]$$

$$\cdot \exp[\frac{D}{N(N+K\cdot D)}\,2\sum_{\ell=1}^{K-1} A_\ell^1 \cos(\ell n\omega_0\tau_n + B_\ell^1)] \tag{6.37}$$

where

$$\begin{aligned} A_\ell^0 \cos B_\ell^0 &= \mathrm{Re}[C_\ell] \\ A_\ell^0 \sin B_\ell^0 &= \mathrm{Im}[C_\ell] \end{aligned} \tag{6.38}$$

$$\begin{aligned} A_\ell^1 \cos B_\ell^1 &= \mathrm{Re}[C_\ell] - \mathrm{Re}[D_\ell] - \mathrm{Re}[E_\ell] + \mathrm{Re}[F_\ell] \\ A_\ell^1 \sin B_\ell^1 &= \mathrm{Im}[C_\ell] - \mathrm{Im}[D_\ell] + \mathrm{Im}[E_\ell] + \mathrm{Im}[F_\ell] \end{aligned} \tag{6.39}$$

and

$$C_\ell \triangleq \sum_{i=0}^{K-1-\ell} z_i(n)^* z_{i+\ell}(n) \tag{6.40}$$

$$D_\ell \triangleq \sum_{i=0}^{K-1-\ell} z_i(n)^* b_0(n)\exp(-j(i+\ell)n\omega_0\tau_s) \tag{6.41}$$

$$E_\ell \triangleq \sum_{i=\ell}^{K-1} z_i(n)^* b_0(n)\exp(-j(i-\ell)n\omega_0\tau_s) \tag{6.42}$$

$$F_\ell \triangleq |b_0(n)|^2\,(K-\ell)\exp(-j\ell n\omega_0\tau_s)\ . \tag{6.43}$$

The essential features of processor structure for this problem are illustrated in Figure 6.3. The energy measuring terms in (6.36) and (6.37) are not included since they will cancel in the formation of the likelihood ratio.

The results of this section are summarized in Table 6.1.

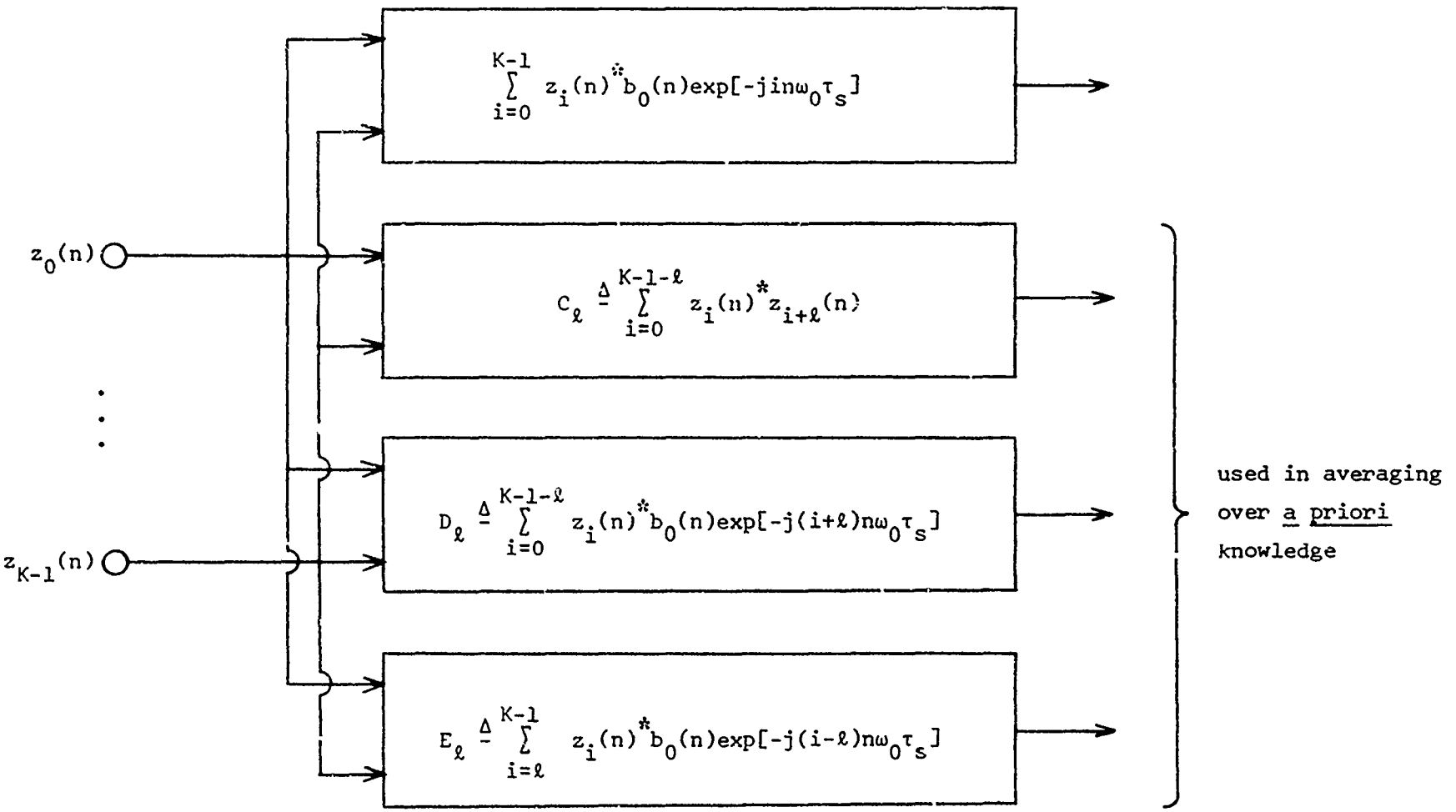


Figure 6.3. Essential Features of the Likelihood Ratio for SKE in NUD.

Conjugate Priors

The single frequency parameter conditional expressions summarized in Table 6.1 can be used directly in calculating the likelihood ratio by simply substituting them in (6.5) and (6.1) with $\Lambda(\underline{z})$ given by the ratio in (6.9). The expressions in Table 6.1 also are the fundamental components required by the general array processor structure implied by (5.12) and illustrated in Figure 5.2. Each sequential/conjugate block is an independent time sequential processor for a single frequency using its own natural conjugate prior. Conjugate priors are used since they are reproducing and thus lend a certain amount of mathematical tractability to the problem.

Recalling from Chapter II, the existence of a sufficient statistic for $\underline{\theta}$ implies the existence of a natural conjugate class of probability density functions under $p(\underline{z}(n)|\underline{\theta})$. From Theorem 2.1, $\underline{\delta}(\underline{z}(n))$ is sufficient for $\underline{\theta}$ if there exists

(1) a function $g[\underline{\delta}(\underline{z}(n)),\underline{\theta}]$ which depends on the observation only through $\underline{\delta}(\cdot)$, and

(2) a function $G(\underline{z}(n))$ which does not depend on $\underline{\theta}$, such that

$$p(\underline{z}(n)|\underline{\theta}) = g[\underline{\delta}(\underline{z}(n)),\underline{\theta}]\, G(\underline{z}(n)) \; . \tag{6.44}$$

The parameter conditional expressions in Table 6.1 suggest the following general form of $\underline{\delta}$

$$\underline{\delta} = [A_0,B_0,\ldots,A_\ell,B_\ell,\ldots,A_{K-1},B_{K-1}]^T \tag{6.45}$$

and of $g[\underline{\delta},\underline{\theta}]$

$$g[\underline{\delta},\underline{\theta}] = \exp[C \sum_{\ell=1}^{K-1} A_\ell \cos(\ell n\omega_0\tau + B_\ell)] \tag{6.4}$$

where C is a constant and τ represents τ_s or τ_n depending on the particular problem being considered. We see from Theorem 2.2 that (6.46) also defines a reproducing class of densities under $p(\underline{z}(n)|\underline{\theta})$. Thus, conjugate priors for these problems have the form

$$p'(\omega_0\tau) = \frac{\exp\left[\sum_{\ell=1}^{K-1} U_\ell \cos(\ell n\omega_0\tau \pm V_\ell)\right]}{\int_{\omega_0\tau} (\text{numerator})\, d\omega_0\tau} \qquad (6.47)$$

where $U_\ell \geq 0$, $|V_\ell| \leq \pi$, and τ again represents τ_s or τ_n .

Each sequential/conjugate block in Figure 5.2 forms a single frequency marginal density based on its particular conjugate prior. The integral required to average over $p'(\omega_0\tau)$ has a form similar to that in the denominator of (6.47). In general, this integral cannot be completed in closed form and some numerical technique must be used for its solution. Note the following special case

$$\frac{1}{2\pi}\int_{-\pi}^{\pi} \exp[A\cos(\omega_0\tau + B)]d\omega_0\tau = I_0(A) \qquad (6.48)$$

where $A \geq 0$, $|B| \leq \pi$, and $I_0(\cdot)$ is the modified Bessel function of order zero.

Single Frequency Parameter Conditional Expressions

A. Gaussian Signal of Uncertain Direction in Gaussian Noise (GUD)

$$\Lambda(\underline{z}(n)|\omega_0\tau_s) = \frac{N}{K\cdot D+N} \exp\left[\frac{S}{N(N+K\cdot S)} \sum_{i=0}^{K-1} z_i(n)^* z_i(n)\right]$$

$$\cdot \exp\left[\frac{S}{N(N+K\cdot S)} 2\left\{\sum_{\ell=1}^{K-1} A_\ell \cos(\ell n\omega_0\tau_s + B_\ell)\right\}\right]$$

where

$$A_\ell \cos B_\ell = \mathrm{Re}[C_\ell]$$

$$A_\ell \sin B_\ell = \mathrm{Im}[C_\ell]$$

and

$$C_\ell \triangleq \sum_{i=0}^{K-1-\ell} z_i(n)^* z_{i+\ell}(n) \ .$$

B. Signal Known Except for Direction in Noise with an Additive Directional Noise Component of Known Direction (SKED in NKD)

$$\Lambda(\underline{z}(n)|\omega_0\tau_s) = \exp\left[-\frac{K[N+(K-1)D]}{N(N+K\cdot D)} |b_0(n)|^2\right.$$

$$\left. + 2\,\mathrm{Re}\left\{\frac{1}{N} z_0(n)^* b_0(n) - \frac{D}{N(N+K\cdot D)} \sum_{i=0}^{K-1} z_i(n)^* b_0(n) \exp(-jin\omega_0\tau_n)\right\}\right]$$

$$\cdot \exp\left[2 \sum_{\ell=1}^{K-1} A_\ell \cos(\ell n\omega_0\tau_s + B_\ell)\right]$$

where

$$A_\ell \cos B_\ell = \mathrm{Re}[G_\ell] - \mathrm{Re}[H_\ell] + \mathrm{Re}[I_\ell]$$

$$A_\ell \sin B_\ell = -\mathrm{Im}[G_\ell] + \mathrm{Im}[H_\ell] - \mathrm{Im}[I_\ell]$$

and

$$G_\ell \triangleq \frac{1}{N} z_\ell(n)^* b_0(n)$$

$$H_\ell \triangleq \frac{D}{N(N+K\cdot D)} \sum_{i=0}^{K-1} z_i(n)^* b_0(n) \exp(-j(i-\ell)n\omega_0\tau_n)$$

Table 6.1

$$I_\ell \triangleq \frac{D}{N(N+K\cdot D)} |b_0(n)|^2 \ (K-\ell) \exp(j\ell n\omega_0\tau_n) \ .$$

C. Signal Known Exactly in Noise with an Additive Directional Noise Component of Uncertain Direction (SKE in NUD)

$$p(\underline{z}(n)|\omega_0\tau_n, H_0) = \frac{1}{\pi^K N^{K-1}(K\cdot D+N)} \exp[-\frac{[N+(K-1)D]}{N(N+K\cdot D)} \sum_{i=0}^{K-1} z_i(\quad (n)]$$

$$\cdot \exp[\frac{D}{N(N+K\cdot D)} 2 \sum_{\ell=1}^{K-1} A_\ell^0 \cos(\ell n\omega_0\tau_n + B_\ell^0)]$$

and

$$p(\underline{z}(n)|\omega_0\tau_n, H_1) = \frac{1}{\pi^K N^{K-1}(K\cdot D+N)} \exp[-\frac{[N+(K-1)D]}{N(N+K\cdot D)} \sum_{i=0}^{K-1} z_i(n)^* z_i(n)$$

$$+ 2 \operatorname{Re}\{\frac{N+(K-1)D}{N(N+K\cdot D)} \sum_{i=0}^{K-1} z_i(n)^* b_0(n) \exp(-jin\omega_0\tau_s)\}$$

$$- \frac{K[N+(K-1)D]}{N(N+K\cdot D)} |b_0(n)|^2] \cdot \exp[\frac{D}{N(N+K\cdot D)} 2 \sum_{\ell=1}^{K-1} A_\ell^1 \cos(\ell n\omega_0\tau_n + B_\ell^1)]$$

where

$$A_\ell^0 \cos B_\ell^0 = \operatorname{Re}[C_\ell]$$

$$A_\ell^0 \sin B_\ell^0 = \operatorname{Im}[C_\ell]$$

$$A_\ell^1 \cos B_\ell^1 = \operatorname{Re}[C_\ell] - \operatorname{Re}[D_\ell] - \operatorname{Re}[E_\ell] + \operatorname{Re}[F_\ell]$$

$$A_\ell^1 \cos B_\ell^1 = \operatorname{Im}[C_\ell] - \operatorname{Im}[\ _\ell] + \operatorname{Im}[E_\ell] + \operatorname{Im}[F_\ell]$$

and

$$C_\ell \triangleq \sum_{i=0}^{K-1-\ell} z_i(n)^* z_{i+\ell}(n)$$

$$D_\ell \triangleq \sum_{i=0}^{K-1-\ell} z_i(n)^* b_0(n) \exp(-j(i+\ell)n\omega_0\tau_s)$$

$$E_\ell \triangleq \sum_{i=\ell}^{K-1} z_i(n)^* b_0(n) \exp(-j(i-\ell)n\omega_0\tau_s)$$

$$F_\ell \triangleq |b_0(n)|^2 (K-\ell) \exp(-j\ell n\omega_0\tau_s) \ .$$

Table 6.1 (continued)

Chapter VII

GENERALIZED PERFORMANCE

The performance of a Bayes optimal processor is summarized by its ROC curve. In the general case where uncertain parameters exist under each hypothesis, knowledge about them is explicitly noted at the outset by an *a priori* probability density function conditonal to H_1 and one conditional to H_0. The processor's performance then becomes a function of their detailed shape. Often, the functional form of these densities is chosen so that various levels of uncertainty are easily modeled and a family of ROC's is reported. The question then arises: What performance would have been achieved under different prior knowledge assumptions (particularly when outside the class of densities modeled)? Or, more deeply: Does some algorithm exist which will operate on a known ROC for a given pair of priors to yield the ROC for a new set of priors? And, if not: Does a canonical intermediate step exist between observation and likelihood ratio statistics which always may be used as a starting point for the calculation of an ROC based on an arbitrary pair of priors? The purpose of this chapter is to pursue these questions. The discussion will use as a basis the fundamental concepts of sufficient statistics and reproducing densities introduced in Chapter II.

The ROC and Observation Statistics

In the evaluation of performance, both a decision rule (i.e., likelihood ratio plus threshold) and ob .vation statistics are needed. Consider

the general case where uncertain parameters exist under each hypothesis. The decision rule is written

$$\Lambda(\underline{R}) = \frac{\int_{\theta_1} p(\underline{R}|\underline{\theta}_1, H_1) p(\underline{\theta}_1)\, d\underline{\theta}_1}{\int_{\theta_0} p(\underline{R}|\underline{\theta}_0, H_0) p(\underline{\theta}_0)\, d\underline{\theta}_0} \underset{D_0}{\overset{D_1}{\gtrless}} \eta \qquad (7.1)$$

where $\underline{R}$ is a vector of observations, $p(\underline{\theta}_1)$ and $p(\underline{\theta}_0)$ are a priori probability density functions of the uncertain parameter vectors $\underline{\theta}_1 \in \theta_1$ and $\underline{\theta}_0 \in \theta_0$, and η is the decision threshold. The likelihood ratio can be viewed simply as a transformation of random variables from the observation space (typically of large dimension) to the one dimensional decision space $[0,\infty)$. It is the conditional distribution of Λ under H_1 and H_0 from which the ROC is calculated (see (2.20) and (2.21)). Peterson, Birdsall, and Fox (Peterson, Birdsall, and Fox, 1954) have shown

$$p(\Lambda|H_1) = \Lambda\, p(\Lambda|H_0) \; . \qquad (7.2)$$

Thus, only the density of Λ conditional to H_0 need actually be obtained.

Owing to the large dimensionality reduction between observation and decision spaces, the question arises as to what "information" is lost in the mapping. Suppose we are given an ROC and the general equation for $\Lambda(\underline{R})$ as in (7.1) for which it summarizes performance. Now, consider a new pair of priors $p'(\underline{\theta}_1)$ and $p'(\underline{\theta}_0)$. Will the ROC under $p(\underline{\theta}_1)$ and $p(\underline{\theta}_0)$ provide us with sufficient information to calculate the ROC under the new prior knowledge? Birdsall has proven two propositions which relate to this question (Birdsall, 1973)

(1) An ROC curve contains insufficient information to specify the observation statistics that led to it, or even to specify the statistics on a real decision axis that led to it.

(2) An ROC does contain sufficient information to specify the statistics on the logarithm of the likelihood ratio.

The second proposition guarantees that we can obtain the densities $p(\ln\Lambda|H_1)$ and $p(\ln\Lambda|H_0)$ from just the information present on the ROC curve. The first states that these densities are insufficient to specify the observation statistics leading to them. A simple example will illustrate this last point.

Consider the detection of a signal known exactly in white Gaussian noise. Recall from Chapter II that performance for this problem is completely characterized by the detectability index

$$d^2 = \frac{2E}{N_0} \tag{2.22}$$

where E is the received signal energy and $N_0/2$ is the noise power spectrum height. Clearly the observation statistics $p(\underline{R}|H_1)$ and $p(\underline{R}|H_0)$ will depend on the exact shape of the signal and height of the noise power spectrum. Note that the value of d^2 depends only on the ratio of signal energy to spectrum height. Thus, many observation situations can yield the same ROC curve. Correspondingly, the information lost in the dimensionality reduction from observation space to decision space is the detailed nature of the observation statistics.

Since the ROC is insufficient to specify the observation statistics leading to it, the conditional statistics required in (7.1) prove to be an impossible refinement. Thus, simply being given an ROC and the underlying priors $p(\underline{\theta}_1)$ and $p(\underline{\theta}_0)$ will not provide enough information to calculate performance under a new pair of priors $p'(\underline{\theta}_1)$ and $p'(\underline{\theta}_0)$.

The Role of Sufficient Statistics

Admittedly, attempting to calculate performance under a new pair of priors with only the original priors and their ROC curve as a basis is an ambitious goal. But it does lead one through the visualization of $\Lambda(\underline{R})$ as a transformation of random variables and, more importantly, that the direct mapping from observation space to decision space under $\Lambda(\underline{R})$ is "lossy."

Consider a new situation. Suppose the performance of a particular detection receiver is desired and assume there are uncertain parameters under each hypothesis. In this case, the conditional statistics in (7.1) will be available explicitly (i.e., their functional form). Once again, view $\Lambda(\underline{R})$ as a transformation of random variables. To compute performance for a particular pair of priors $p(\underline{\theta}_1)$ and $p(\underline{\theta}_0)$, the transformation through Λ is completed to yield $p(\Lambda|H_1)$ and $p(\Lambda|H_0)$ from which the ROC can be determined. Now, if the performance for an entirely different pair of priors $p'(\underline{\theta}_1)$ and $p'(\underline{\theta}_0)$ is desired, the transformation must begin again with the observation statistics and proceed as outlined above. The question arises: Is there some well defined intermediate step between the observation statistics and those of Λ which we might use as a basis (i.e., a new "observation space" of lower dimensionality than $\mathcal{R}$) for the computation of performance for any pair of priors?

The remainder of the discussion will assume that $p(\underline{R}|\underline{\theta}_1,H_1)$ and $p'(\underline{R}|\underline{\theta}_0,H_0)$ admit sufficient statistics of finite dimension for $\underline{\theta}_1$ and $\underline{\theta}_0$, respectively. The expression in (7.1) becomes

$$\Lambda(\underline{R}) = \frac{G_1(\underline{R})\int_{\Theta_1} g_1[\underline{\delta}_1(\underline{R}),\underline{\theta}_1]\, p(\underline{\theta}_1)\, d\underline{\theta}_1}{G_0(\underline{R})\int_{\Theta_0} g_0[\underline{\delta}_0(\underline{R}),\underline{\theta}_0]\, p(\underline{\theta}_0)\, d\underline{\theta}_0} \underset{D_0}{\overset{D_1}{\gtrless}} \eta \qquad (7.3)$$

where $\underline{\delta}_1(\underline{R})$ and $\underline{\delta}_0(\underline{R})$ are the finite dimensional sufficient statistics as in (2.12). It appears from (7.3) that a plausible intermediate step would be the space formed by the random variables $G_1(\underline{R})/G_0(\underline{R})$, $\underline{\delta}_1(\underline{R})$, and $\underline{\delta}_1(\underline{R})$. This step occurs just prior to the averaging over $p(\underline{\theta}_1)$ and $p(\underline{\theta}_0)$ and is illustrated in Figure 7.1. Recall that only the transformation conditional to H_0 need be carried through due to the formula in (7.2). Two observations can be made

(1) Uncertain parameters under H_1 only.

In this special case of the general problem, the observation statistics conditional to H_0 stay the same wher new prior knowledge is assumed. Thus, the proposed intermediate step is a valid point from which calculations of perforance for any prior knowledge can be started.

(2) Uncertain parameters under H_1 and H_0.

Since the observation statistics under H_0 essentially have $p(\underline{\theta}_0)$ embedded in them, the proposed intermediate step will be valuable only when $p(\underline{\theta}_0)$ remains fixed and $p(\underline{\theta}_1)$ alone is allowed to change.

An Example: SKEP

As a specific illustration of the natural intermediate step discussed in the previous section, consider the problem of detecting a signal known except for phase (SKEP) in additive white Gaussian noise bandlimited to W Hz. (Roberts, 1965). Since there is a single uncertain parameter and it exists under the H_1 hypothesis only, the likelihood ratio can be written as

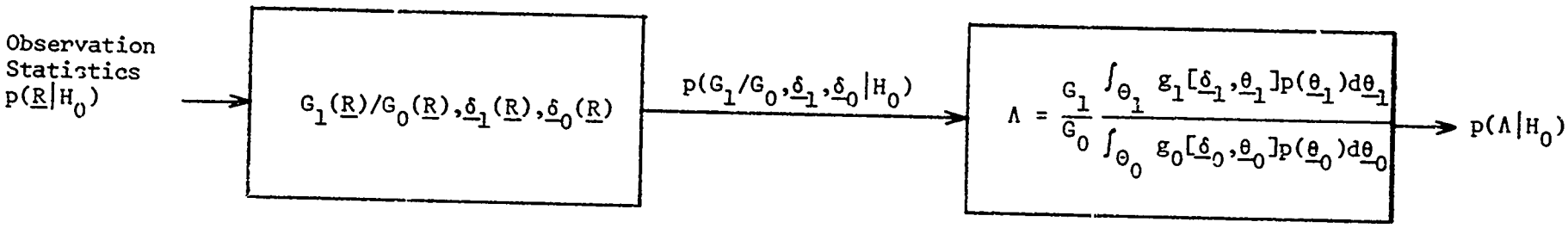


Figure 7.1. Two Step Sequence for the Calculation of Performance.

$$\Lambda(\underline{R}) = \exp[-d^2/2]\int_{\phi} \exp[d(x(\underline{R})\cos\phi + y(\underline{R})\sin\phi)]p(\phi)d\phi \underset{D_0}{\overset{D_1}{\gtrless}} n \qquad (7.4)$$

where $$d^2 = \frac{2E}{N_0}$$

$$E = \frac{1}{2W}\sum_{i=1}^{2WT} s^2(t_i) = \frac{1}{2W}\sum_{i=1}^{2WT} [m(t_i)\cos(\omega t_i - \phi)]^2$$

$N_0/2$ = height of the noise power spectrum

$$x(\underline{R}) = \frac{1}{W\sqrt{2EN_0}}\sum_{i=1}^{2WT} R(t_i)m(t_i)\cos\,\omega t_i$$

and $$y(\underline{R}) = \frac{1}{W\sqrt{2EN_0}}\sum_{i=1}^{2WT} R(t_i)m(t_i)\sin\,\omega t_i \cdot$$

In the notation of sufficient statistics, (7.4) becomes

$$\Lambda(\underline{R}) = \int_{\phi} g[\underline{\delta}(\underline{R}),\phi]p(\phi)d\phi \underset{D_0}{\overset{D_1}{\gtrless}} n \qquad (7.5)$$

where $$g[\underline{\delta}(\underline{R}),\phi] = \exp[-d^2/2 + d(\delta_{01}(\underline{R})\cos\phi + \delta_{02}(\underline{R})\sin\phi)]$$

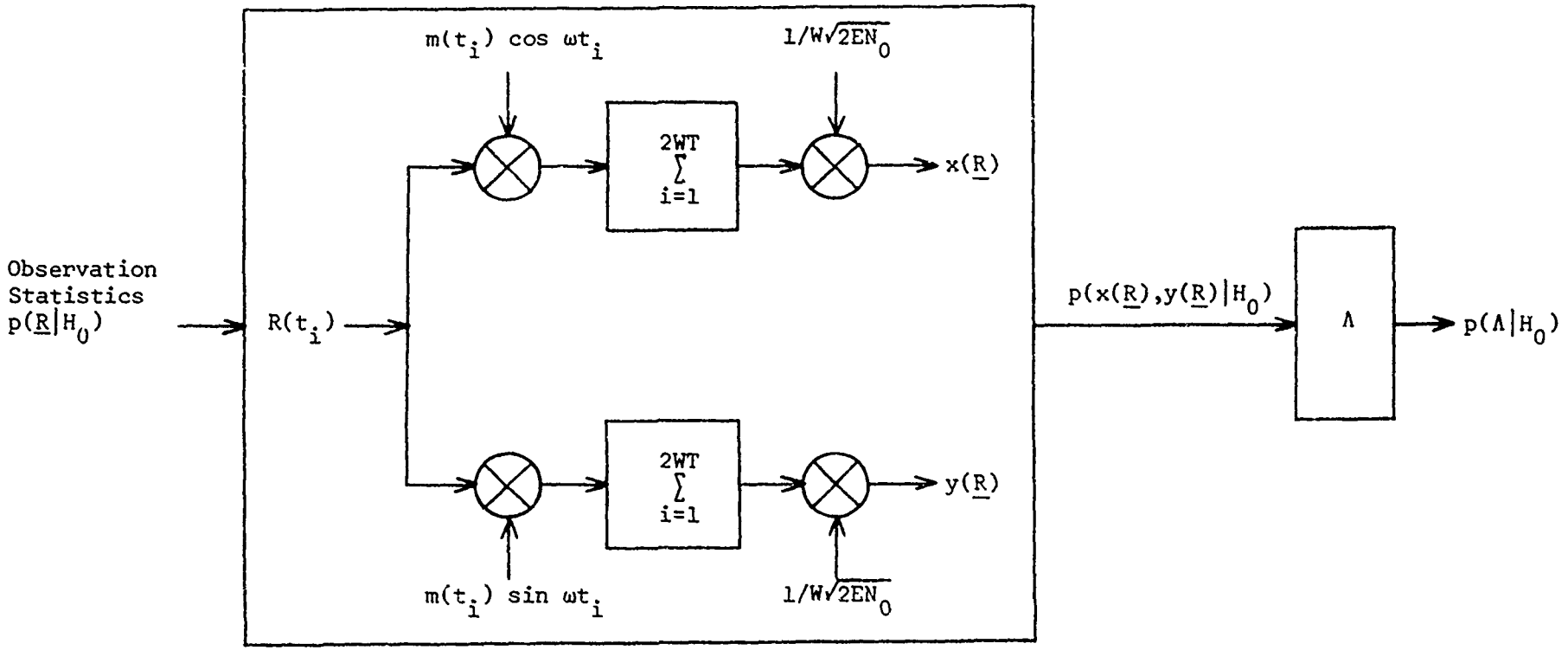
and $$\underline{\delta}(\underline{R}) = [\delta_{01}(\underline{R}),\delta_{02}(\underline{R})]^T = [x(\underline{R}),y(\underline{R})]^T .$$

Conditional to H_0, $x(\underline{R})$ and $y(\underline{R})$ are independent Gaussian random variables with zero mean and unit variance. Thus, the desired intermediate density as indicated in Figure 7.2 is

$$p(\underline{\delta}|H_0) = p(x(\underline{R}),y(\underline{R})|H_0)$$

$$= N(0,1) \cdot N(0,1) \qquad (7.6)$$

where $N(0,1)$ denotes a Gaussian density with zero mean and unit variance.



$$\Lambda = \exp[-d^2/2] \int_{\phi} \exp[d\cdot(x(\underline{R})\cos\phi+y(\underline{R})\sin\phi)]p(\phi)d\phi$$

Figure 7.2. Two Step Sequence for SKEP.

We may now view $p(\underline{\delta}|H_0)$ as new "observation statistics" and performance for any prior $p(\phi)$ can be obtained by completing the indicated transformation of random variables. Note that a significant dimensionality reduction has taken place without loosing the ability to compute performance for an arbitrary prior.

When a priori knowledge is chosen from the class of densities

$$p(\phi) = \frac{\exp[A_0 \cos (B_0-\phi)]}{2\pi\, I_0(A_0)} \quad , \; 0 \le \phi \le 2\pi \qquad (7.7)$$

$$= 0 \quad , \text{ otherwise}$$

Roberts has shown that the likelihood ratio becomes (Roberts, 1965)

$$\Lambda = \frac{\exp[-d^2/2]\; I_0(A_1)}{I_0(A_0)} \qquad (7.8)$$

where $A_1^2 = [d\cdot x(\underline{R}) + A_0 \cos B_0]^2 + [d\cdot y(\underline{R}) + A_0 \sin B_0]^2$.

The processor block diagram is shown in Figure 7.3. Note that the sufficient statistic $\underline{\delta}(\underline{R}) = [x(\underline{R}),y(\underline{R})]^T$ is calculated just prior to the incorporation of the a priori knowledge.

Performance: Optimal and Suboptimal

The observations at the end of the second section pointed out the usefulness of the joint density $p(G_1/G_0,\underline{\delta}_1,\underline{\delta}_0|H_0)$ as an intermediate step in the calculation of performance for an arbitrary prior. In general, this step will enable us to calculate optimal performance only when $p(\underline{\theta}_1)$ is allowed to be arbitrary.

The intermediate step potentially may be beneficial in a different sense when both $p(\underline{\theta}_1)$ and $p(\underline{\theta}_0)$ are allowed to be arbitrary. Although optimal performance cannot be obtained, the processor's sensitivity to

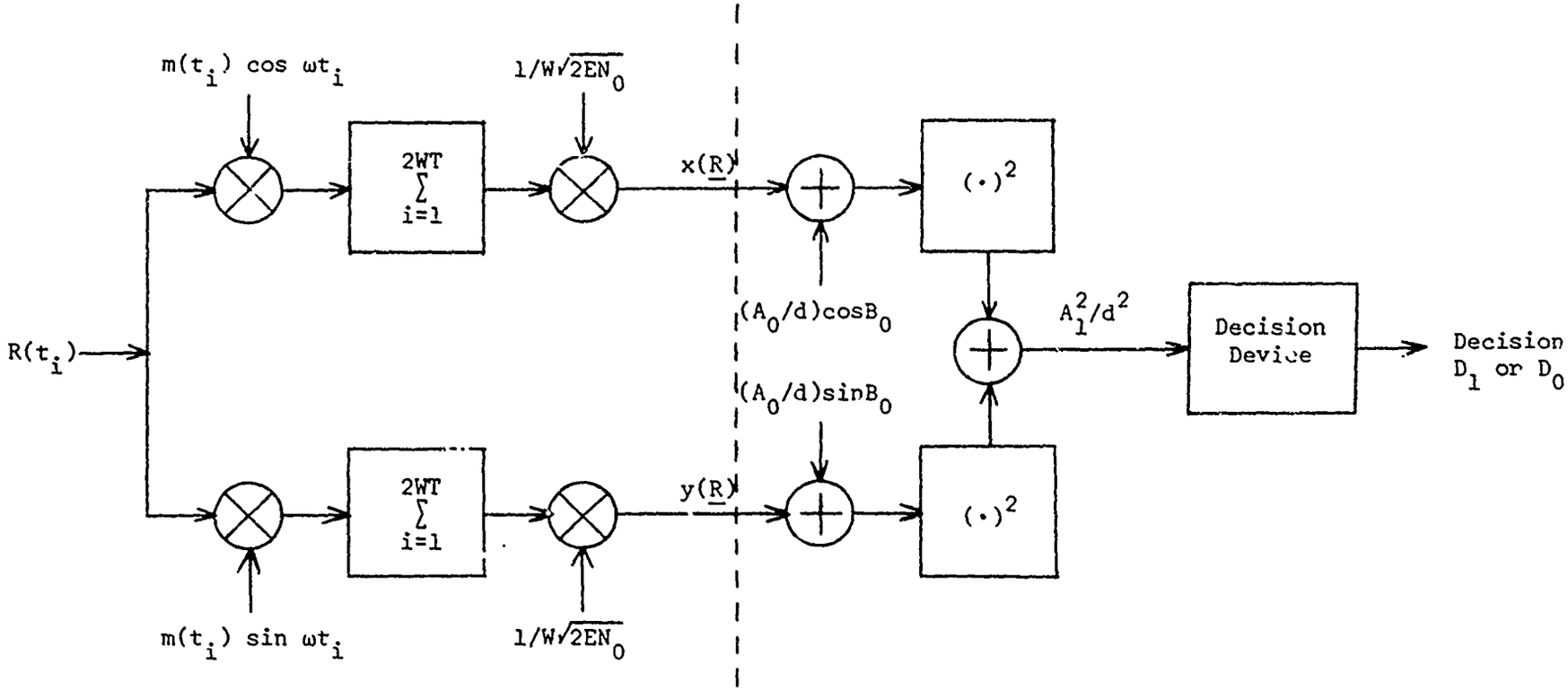


Figure 7.3. SKEP Processor. (Roberts, 1965)

an accurate incorporation of a priori knowledge can be studied. In this case, the intermediate density conditional to both H_0 and H_1 will be needed since the secondary processor they feed is no longer optimal (i.e., the observation statistics do not have embedded in them the priors utilized in the processor). One such problem of potential interest would be where the single uncertain parameter of noise power spectrum height exists under H_1 and H_0.

A major benefit of a natural intermediate step between observation space and the likelihood ratio is found in performance calculation via computer simulation. When only the prior knowledge under H_1 is arbitrary, the intermediate step shown in Figure 7.1 is a valid point from which performance calculations can begin. As in the SKEP problem, this step can represent a significant reduction in dimensionality from that of the observation space. The procedure would be to generate observation vectors $\underline{R}$ conditional to H_0. For each $\underline{R}$, the values of G_1/G_0, $\underline{\delta}_1$, and $\underline{\delta}_0$ would be calculated and retained (in effect, generating a discrete version of $p(G_1/G_0, \underline{\delta}_1, \underline{\delta}_0 | H_0)$). Then, for every $p(\underline{\theta}_1)$ of interest, the new "observation vectors" consisting of $(G_1/G_0, \underline{\delta}_1, \underline{\delta}_0)$ would be used as input to the secondary processor shown in Figure 7.1. The resulting collection of Λ's would be used to form an approximation to $p(\Lambda | H_0)$ and performance calculated.

Chapter VIII

THREE SPECIFIC PROBLEMS AND THEIR PERFORMANCE

Detection performance in terms of the receiver operating characteristic (ROC) curve for the following three specific problems now will be discussed

(1) Gaussian Signal of Uncertain Direction in Gaussian Noise (GUD)

(2) Signal Known Except for Direction in Noise with an Additive Directional Noise Component of Known Direction (SKED in NKD)

(3) Signal Known Exactly in Noise with an Additive Directional Noise Component of Uncertain Direction (SKE in NUD).

Their parameter conditional joint density expressions were derived in Chapter VI and summarized in Table 6.1. To evaluate performance, the three optimal array processors were implemented on the computer. From Monte Carlo simulations of each processor, the distribution of $\Lambda(\underline{z})$ under H_0 was obtained. This distribution then was used to calculate the detection and false alarm probabilities which are summarized on the ROC curve. Several important concepts in the calculation of performance for likelihood ratio processors via computer simmulation are discussed in Appendix F.

Prior Knowledge

In each of the three problems, uncertainty arises in the location of a signal or noise source. The array processor sees location uncertainty reflected in terms of an uncertain time delay of the directional source between adjacent elements. And, in turn, this corresponds to an uncertain phase delay in the frequency domain where the processing is actually carried

out. Thus, a priori knowledge on location will be summarized by the probability density function $p(\omega_0\tau)$ where τ represents either τ_s or τ_n. Such knowledge is explicitly incorporated into the likelihood ratio processor through the averaging in (6.1) and (6.5).

The density $p(\omega_0\tau)$ was chosen from the natural conjugate class of probability density functions for this group of three problems whose form is that of several terms as in (6.47) multiplied together (one for each frequency index). Specifically

$$p(\omega_0\tau) = \frac{N}{2\pi\, I_0(A_0)} \exp[A_0 \cos(N\omega_0\tau + B_0)] \quad , \quad \frac{-\pi}{N} \le \omega_0\tau \le \frac{\pi}{N}$$
$$= 0 \quad , \quad \text{otherwise} \tag{8.1}$$

where $A_0 \ge 0$, $-\pi \le B_0 \le \pi$, and $I_0(\cdot)$ is the modified Bessel function of order zero. For all cases where performance is reported here, $N = 8$ and $B_0 = 0$. As Figure 8.1 indicates, varying the parameter A_0 from zero to infinity models a wide range of uncertainty from diffuse to very precise prior knowledge. The array elements are assumed one half wavelength apart at frequency $N\omega_0/2\pi$ Hz. Thus, (8.1) corresponds to a location uncertainty over ±90° in physical angle from broadside to the array or $\pm\pi$ in phase at frequency $N\omega_0/2\pi$ Hz. The parameters A_0 and B_0 should not be confused with similarly denoted parameters in Table 6.1.

Performance

The remainder of this chapter will be devoted to the discussion and comparison of performance results for the three problems enumerated earlier. Each ROC is labeled by the A_0 value corresponding to a particular level of uncertainty. The array sizes investigated were for three and nine elements as denoted by the parameter K. Both the Gaussian noise which is independent

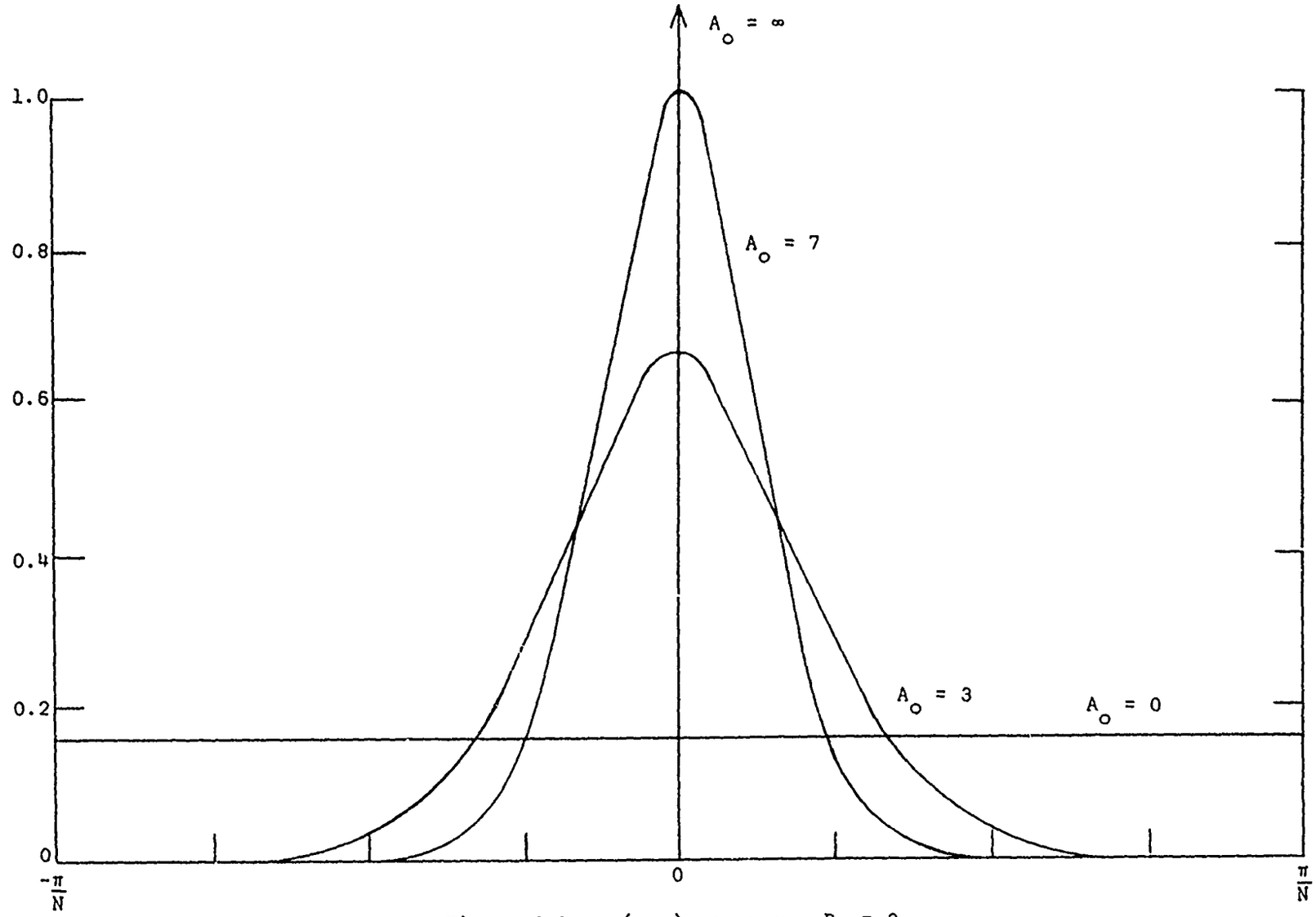


Figure 8.1. $p(\omega_0\tau)$ vs. $\omega_0\tau$; $B_0 = 0$.

from sensor to sensor and the additive directional Gaussian noise have power spectra which are white and bandlimited to $N\omega_0/2\pi$ Hz. Their spectral heights are denoted by N and D, respectively.

GUD

Performance curves for the processor which is to decide between presence or absence of a Gaussian signal of uncertain direction in Gaussian noise which is independent from sensor to sensor are given in Figures 8.2-8.6. The five figures are for a wide band signal source whose power spectrum is white and bandlimited to $N\omega_0/2\pi$ Hz. Its spectral height is denoted by S. The likelihood ratio for this problem is given by

$$\Lambda(\underline{z}) = \int_{\omega_0\tau_s} \prod_{n=0}^{N} \Lambda(\underline{z}(n)|\omega_0\tau_s)p(\omega_0\tau_s)d\omega_0\tau_s \qquad (8.2)$$

where $\Lambda(\underline{z}(n)|\omega_0\tau_s)$ can be found in Table 6.1 A.

Note from Figures 8.2-8.4 that increasing location uncertainty leads to a greater degradation in performance the larger the array size. The expected value of the total signal energy processed in Figures 8.2 and 8.3 has been kept constant to make this comparison more visible.

Figures 8.5 and 8.6 evaluate performance for a particular suboptimal processor which has been placed in the same environment of uncertainty as discussed above. This suboptimal processor has the estimate and plug structure as illustrated in Figure 3.4. Since the uncertain parameter of signal source location exists under H_1 only, $\hat{\underline{\theta}}_1$ is the sole estimate required. In this case, the estimate is not data dependent and is fixed at $\hat{\theta}_1 = (\hat{\omega_0\tau_s}) = 0$. The parameter conditional expression into which $\hat{\underline{\theta}}_1$ is plugged is given by

$$\Lambda(\underline{z}|\underline{\theta}_1) = \prod_{n=0}^{N} \Lambda(\underline{z}(n)|\omega_0\tau_s) \qquad (8.3)$$

where $\Lambda(\underline{z}(n)|\omega_0\tau_s)$ can be found in Table 6.1 A. The resulting processor is a realization of that derived by Bryn (Bryn, 1962) with a front end consisting of a beamformer looking broadside to the array. Note that even in a relatively uncertain environment, the suboptimal three element array processor does not suffer greatly in performance loss. However, the suboptimal nine element array processor suffers a significant drop in performance for all levels of uncertainty less than precise knowledge. These two figures also should be compared to Figures 8.2 and 8.4. Particularly for the nine element array, these results point out the necessity of properly incorporating *a priori* knowledge into the array processor design.

GUD in WKD

Performance curves are given in Figures 8.7-8.18 for the processor which is to decide presence or absence of a signal known except for direction in Gaussian noise consisting of a component which is independent from sensor to sensor and an additive component arising from a source of known direction. The likelihood ratio for this problem is given by

$$\Lambda(\underline{z}) = \int_{\omega_0\tau_s} \prod_{n=0}^{N} \Lambda(\underline{z}(n)|\omega_0\tau_s) p(\omega_0\tau_s)\, d\omega_0\tau_s \qquad (8.4)$$

where $\Lambda(\underline{z}(n)|\omega_0\tau_s)$ can be found in Table 6.1 B. A single frequency signal is assumed at $N\omega_0/2\pi$ Hz. Its energy over the observation interval is given by $E = 2\, b_0(N)^* b_0(N)$. The first five figures are for a noise component of known direction of zero spectral height. This reduces the problem to the known signal counterpart of the GUD problem previously discussed. Performance for this special case of the optimal processor originally was reported by Gallop (Gallop, 1971; Gallop and Nolte, 1974). Those results are not identical to that reported here since Gallop assumed the zeroth array element

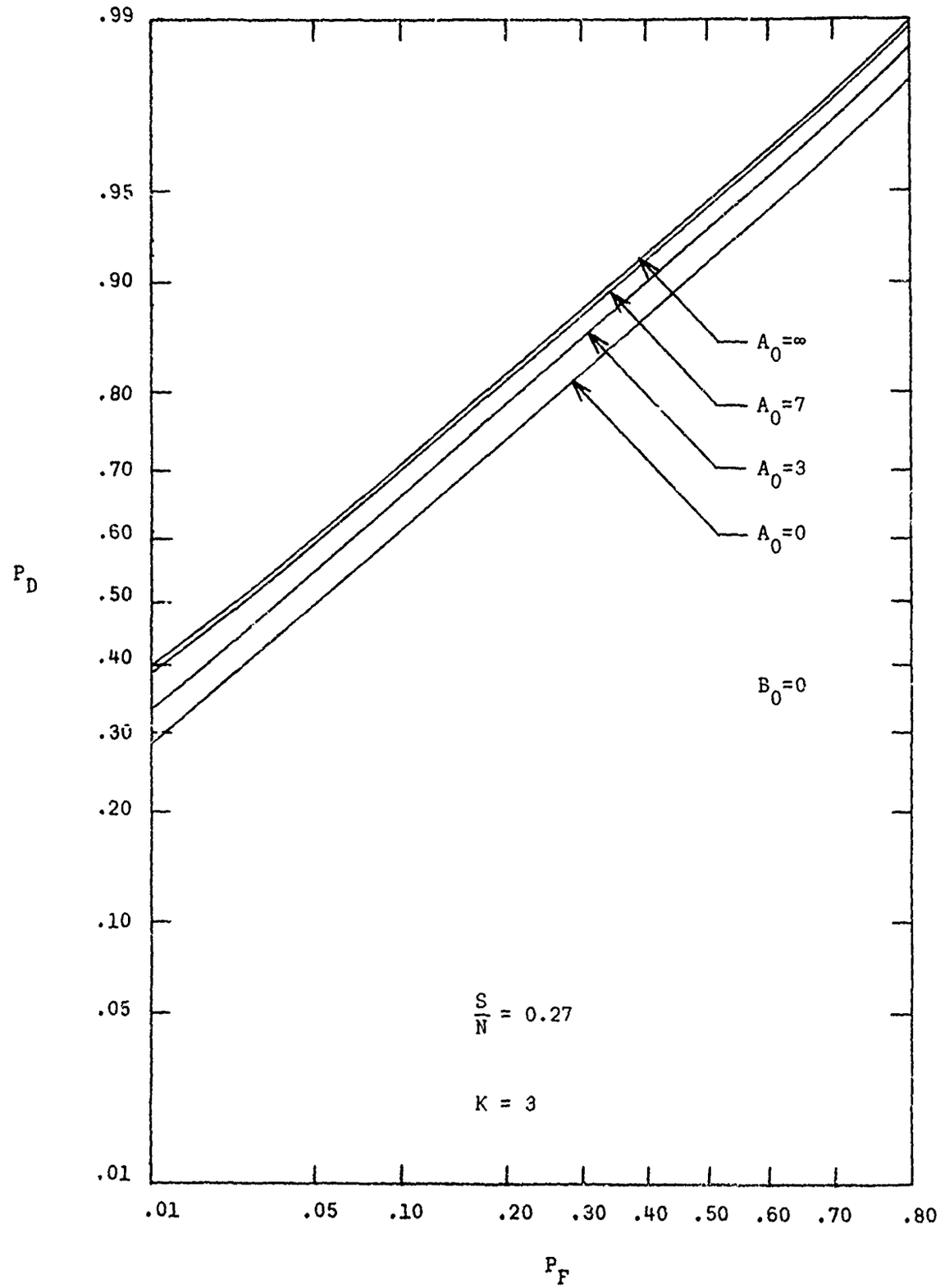


Figure 8.2. Performance of the Optimal GUD Array Processor.

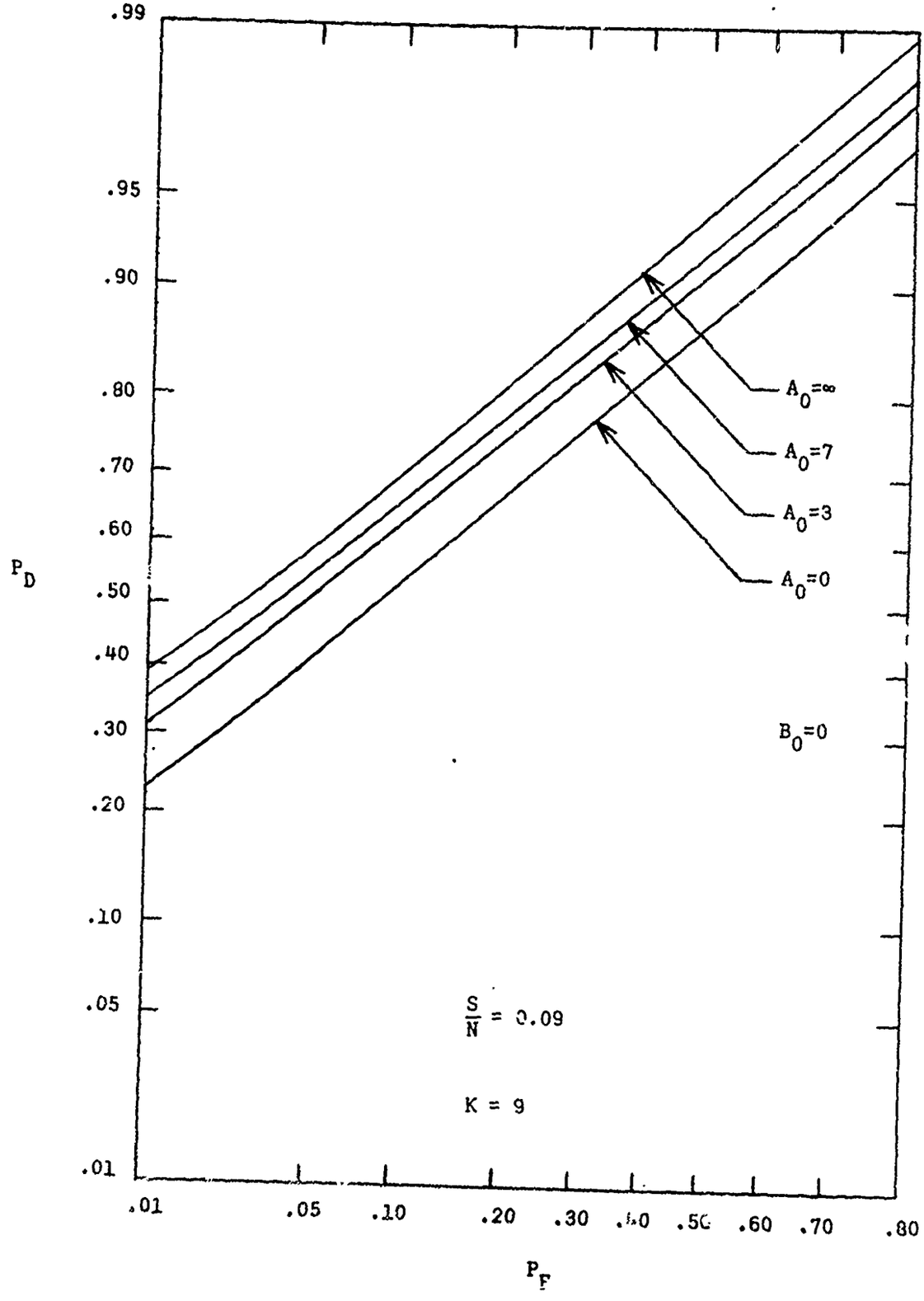


Figure 8.3. Performance of the Optimal GUD Array Processor.

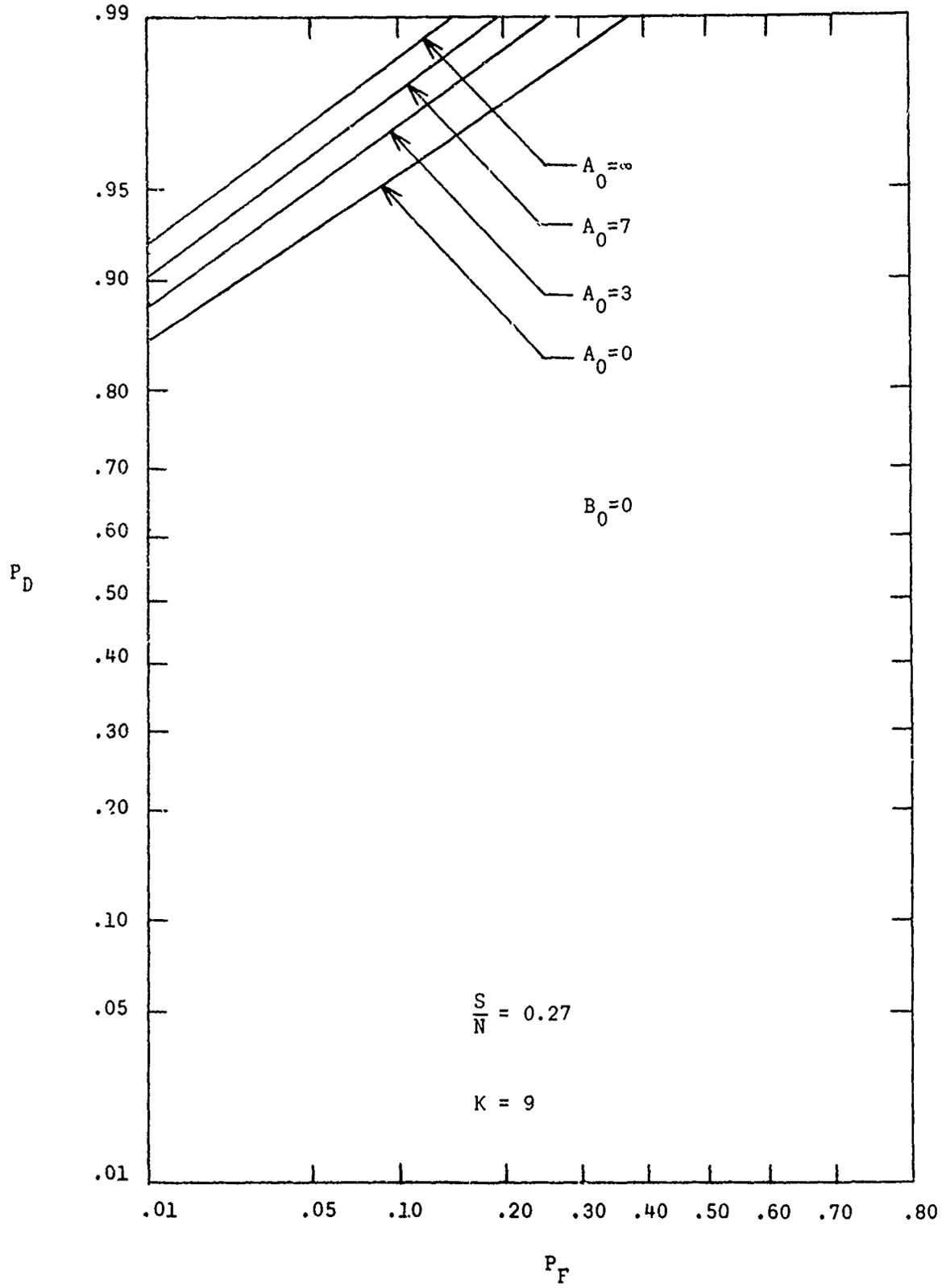


Figure 8.4. Performance of the Optimal GUD Array Processor.

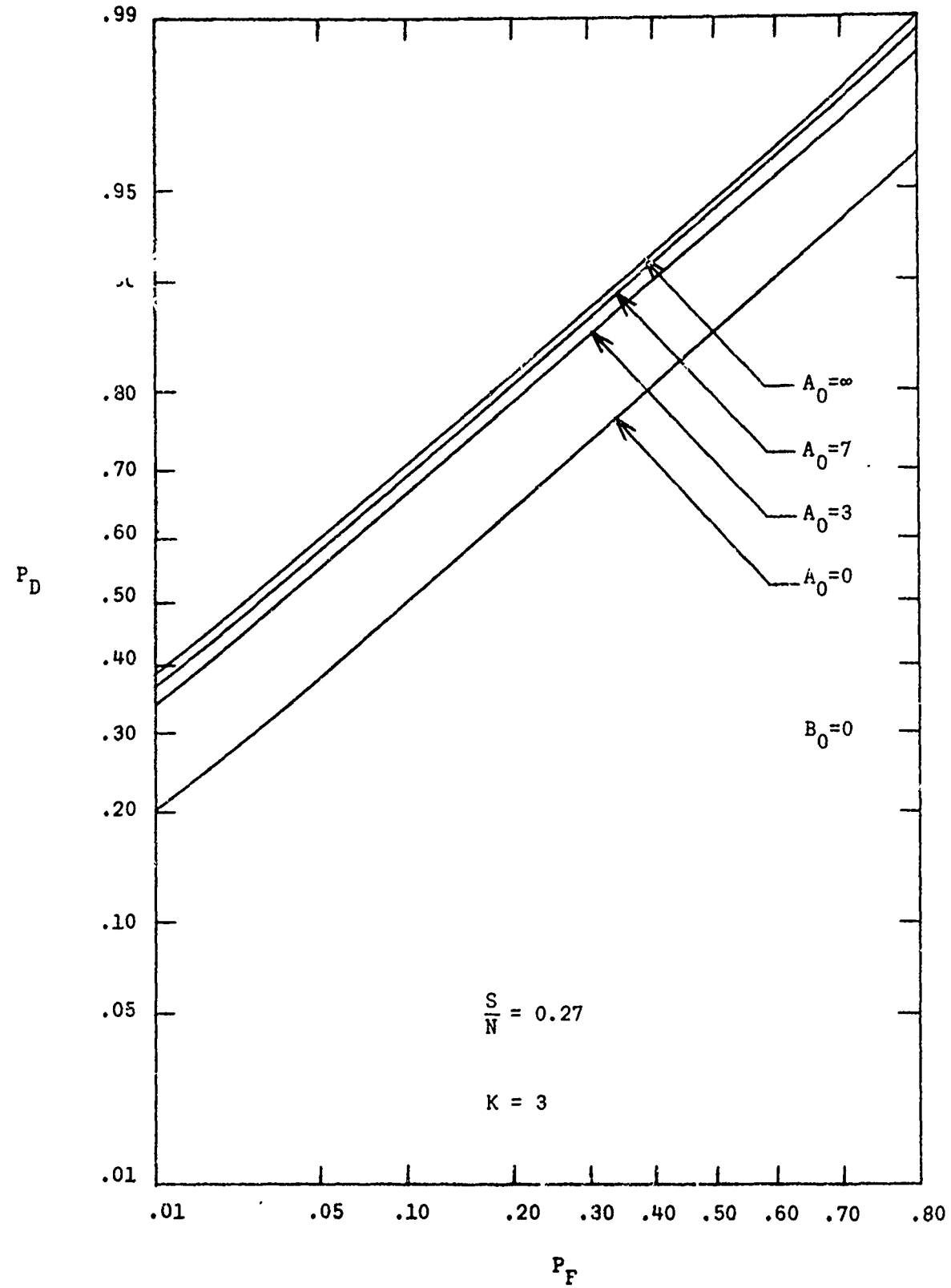


Figure 8.5. Performance of the Suboptimal GUD Array Processor.

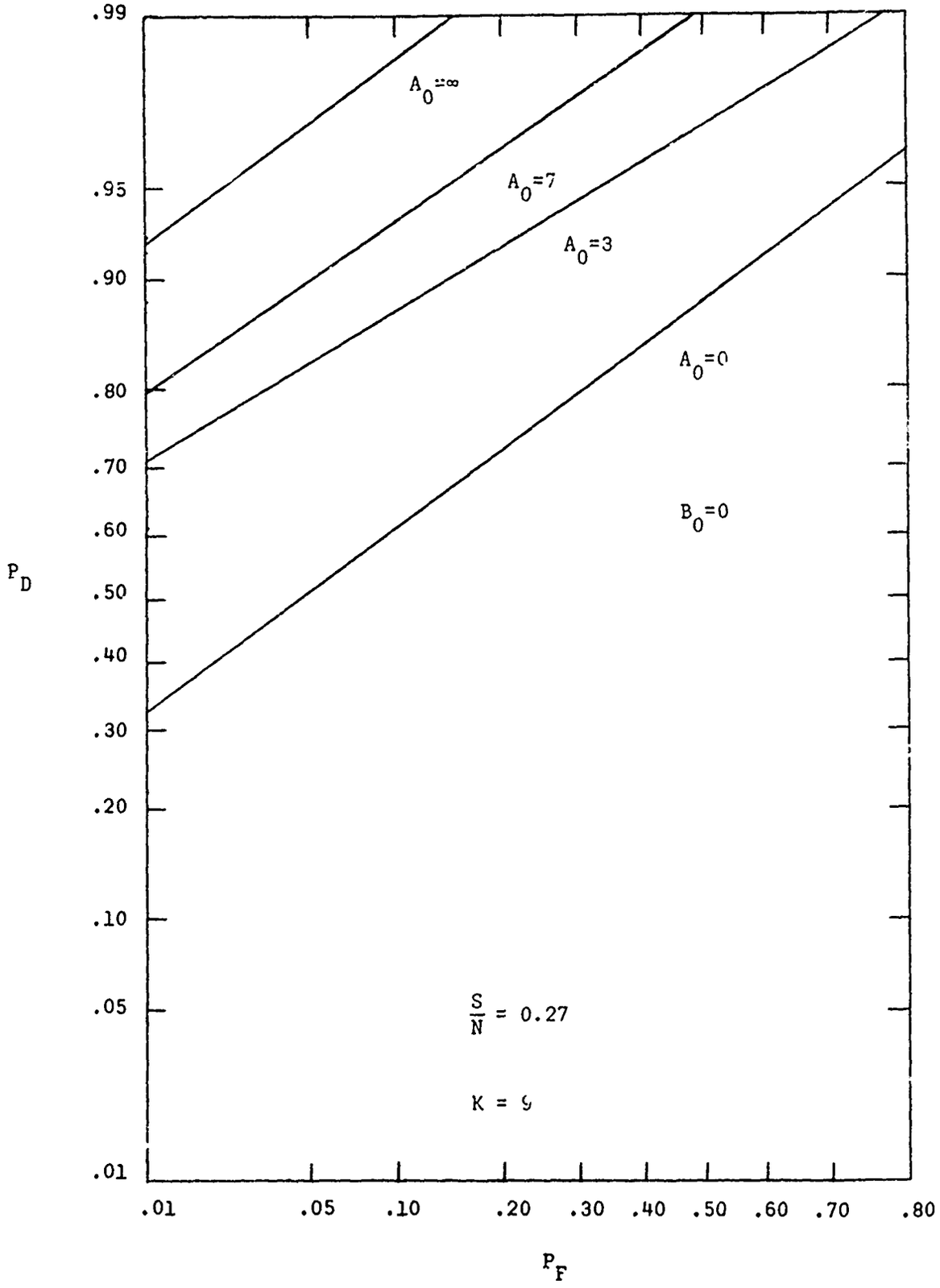


Figure 8.6. Performance of the Suboptimal GUD Array Processor.

was located at the array center instead of the array end as is assumed here. The remaining seven figures are for a noise component of known direction of unit spectral height.

Note from Figures 8.7-8.9 that increasing location uncertainty leads to a greater degradation in performance the larger the array size. This is particularly evident in Figures 8.7 and 8.8 where the total signal energy processed has been held constant.

Figures 8.10 and 8.11 evaluate performance for a particular suboptimal processor which has been placed in the same environment of uncertainty as discussed above. This suboptimal processor has the estimate and plug structure as illustrated in Figure 3.4. Since the uncertain parameter of signal source location exists under H_1 only, $\hat{\underline{\theta}}_1$ is the sole estimate required. In this case, the estimate is not data dependent and is fixed at $\hat{\underline{\theta}}_1 = (\hat{\omega_0\tau_s}) = 0$. The parameter conditional expression into which $\hat{\underline{\theta}}_1$ is plugged is given by

$$\Lambda(\underline{z}|\underline{\theta}_1) = \prod_{n=0}^{N} \Lambda(\underline{z}(n)|\omega_0\tau_s) \tag{8.5}$$

where $\Lambda(\underline{z}(n)|\omega_0\tau_s)$ can be found in Table 6.1 B. The resulting processor is a realization of that derived by Mermoz (Mermoz, 1964; Horton, 1969) with a front end consisting of a beamformer looking broadside to the array. For both the three and nine element processors, significant performance degradation is suffered for all levels of uncertainty other than precise knowledge. This is particularly evident the larger the array. Once again, a comparison of these results with those in Figures 8.7 and 8.9 point out the necessity of properly incorporating a priori knowledge into the array processor design.

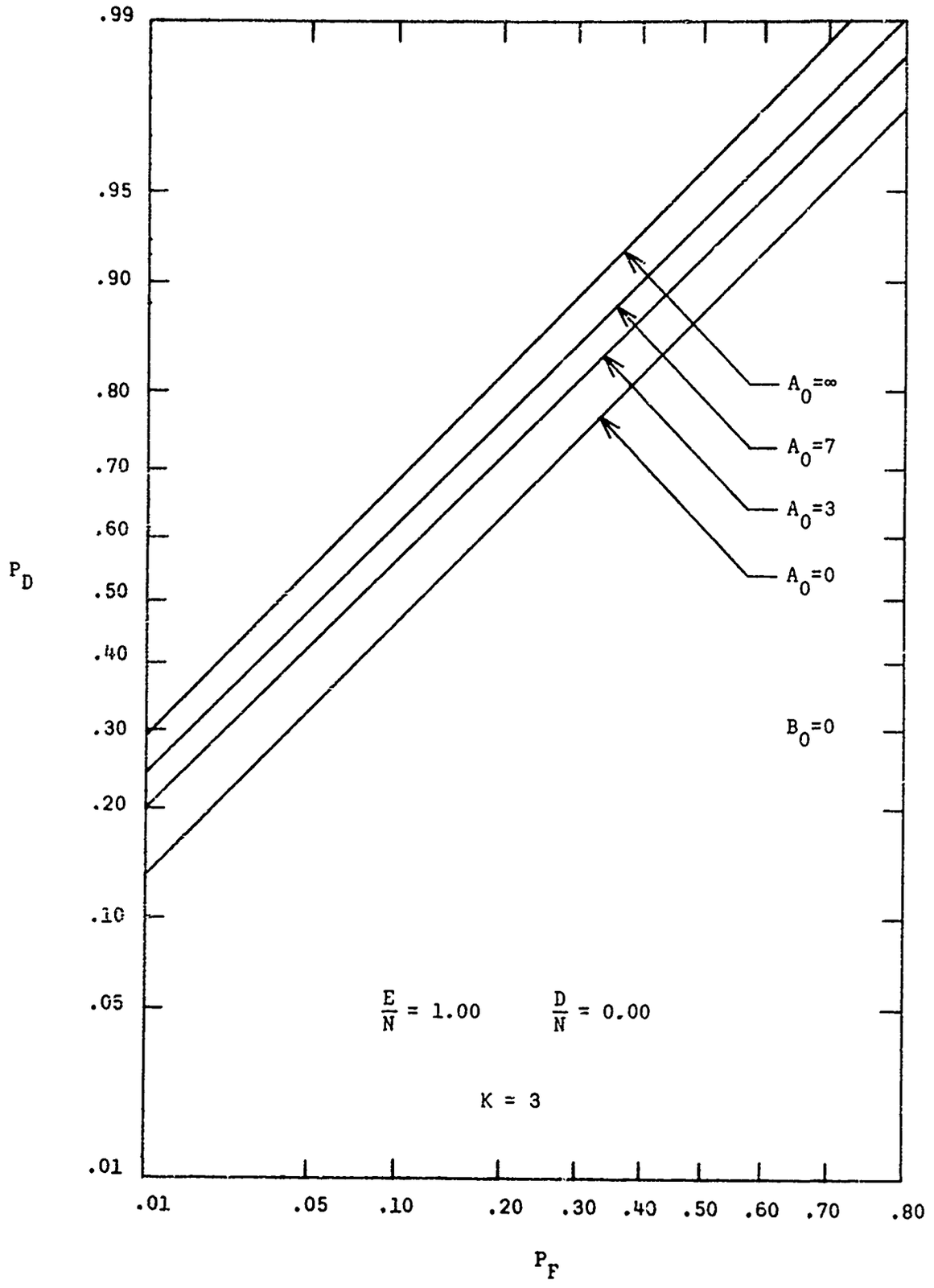


Figure 8.7. Performance of the Optimal SKED in NKD Array Processor.

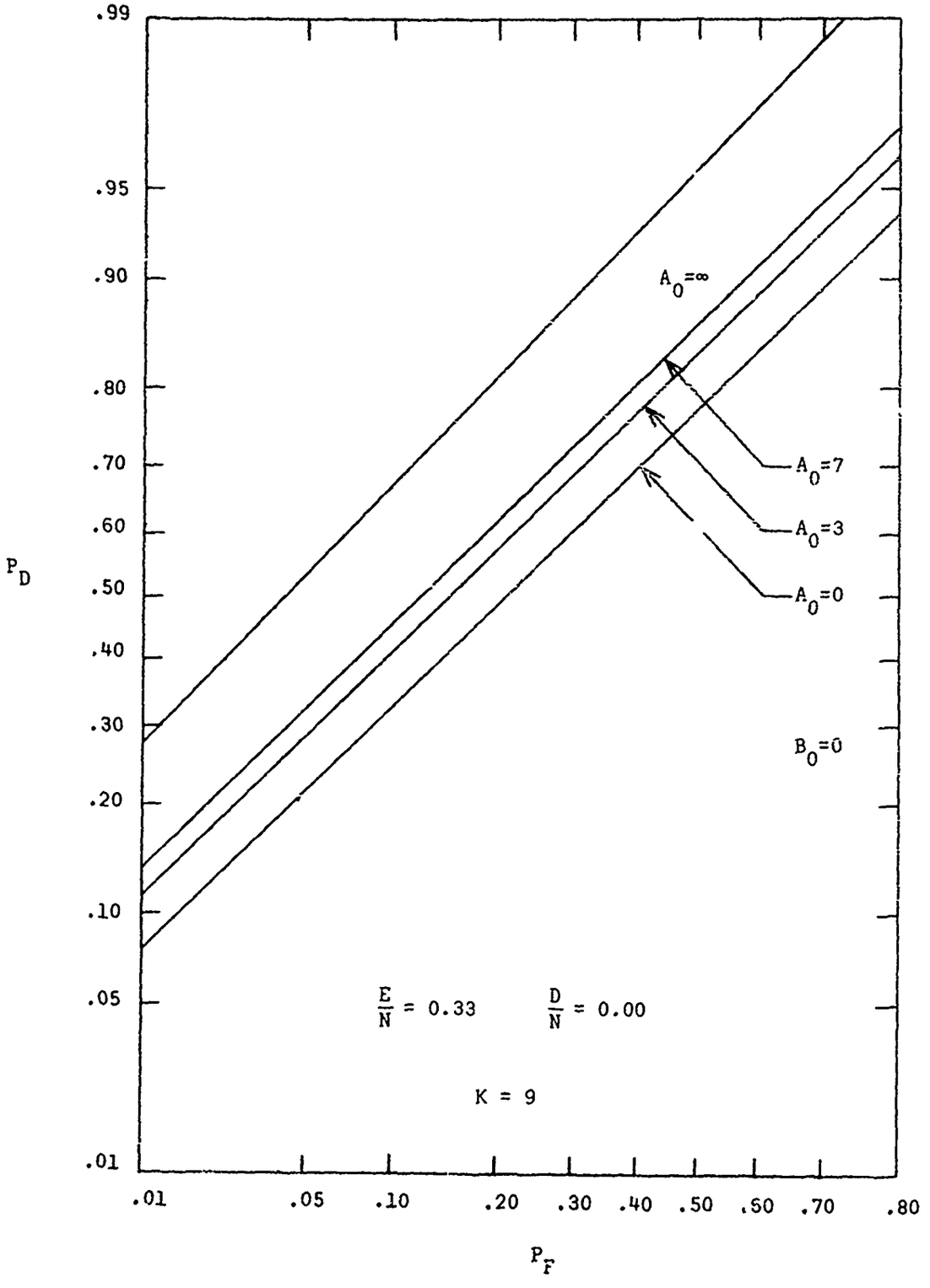


Figure 8.8. Performance of the Optimal SKED in NKD Array Processor.

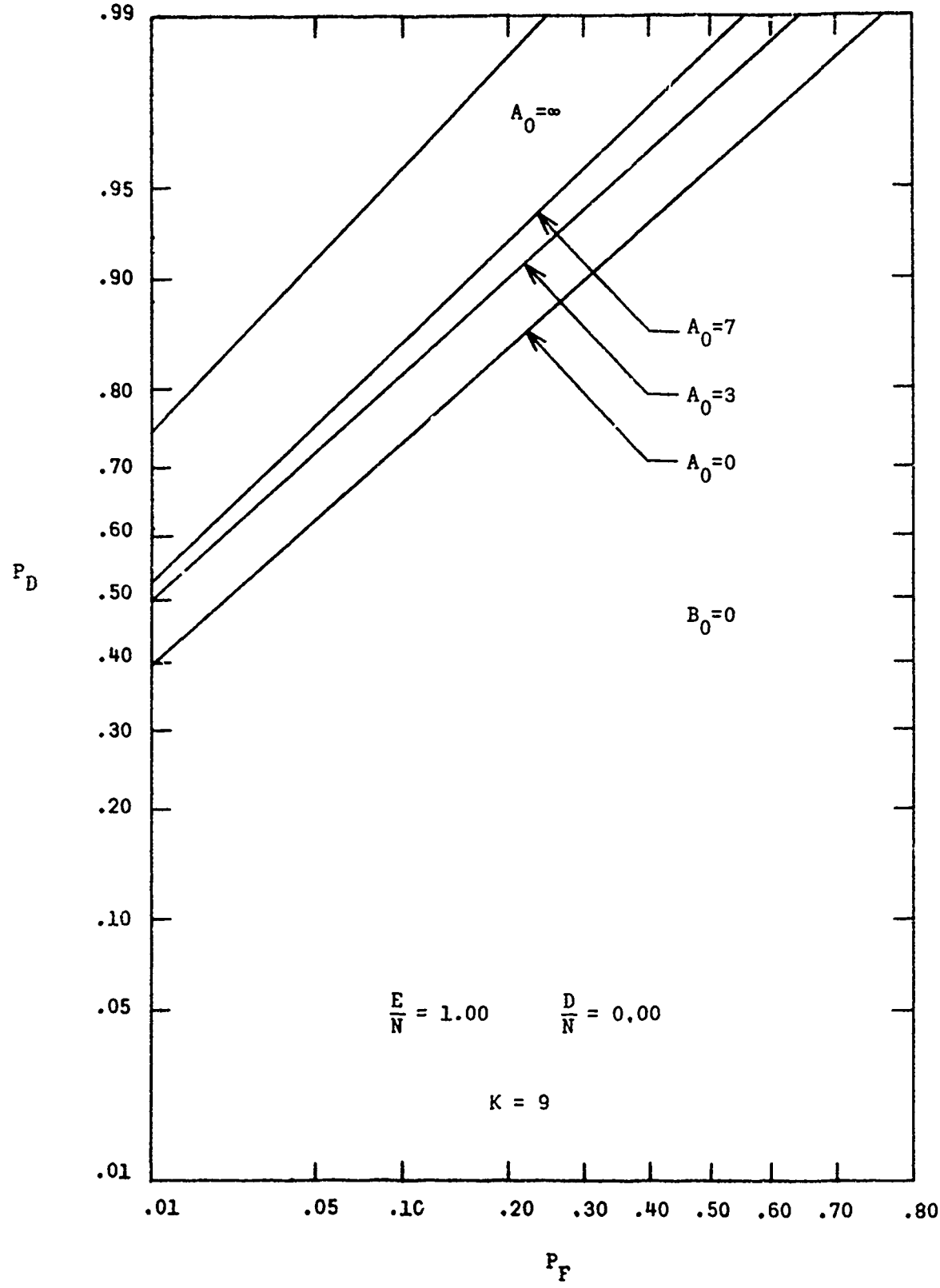


Figure 8.9. Performance of the Optimal SKED in NTD Array Processor.

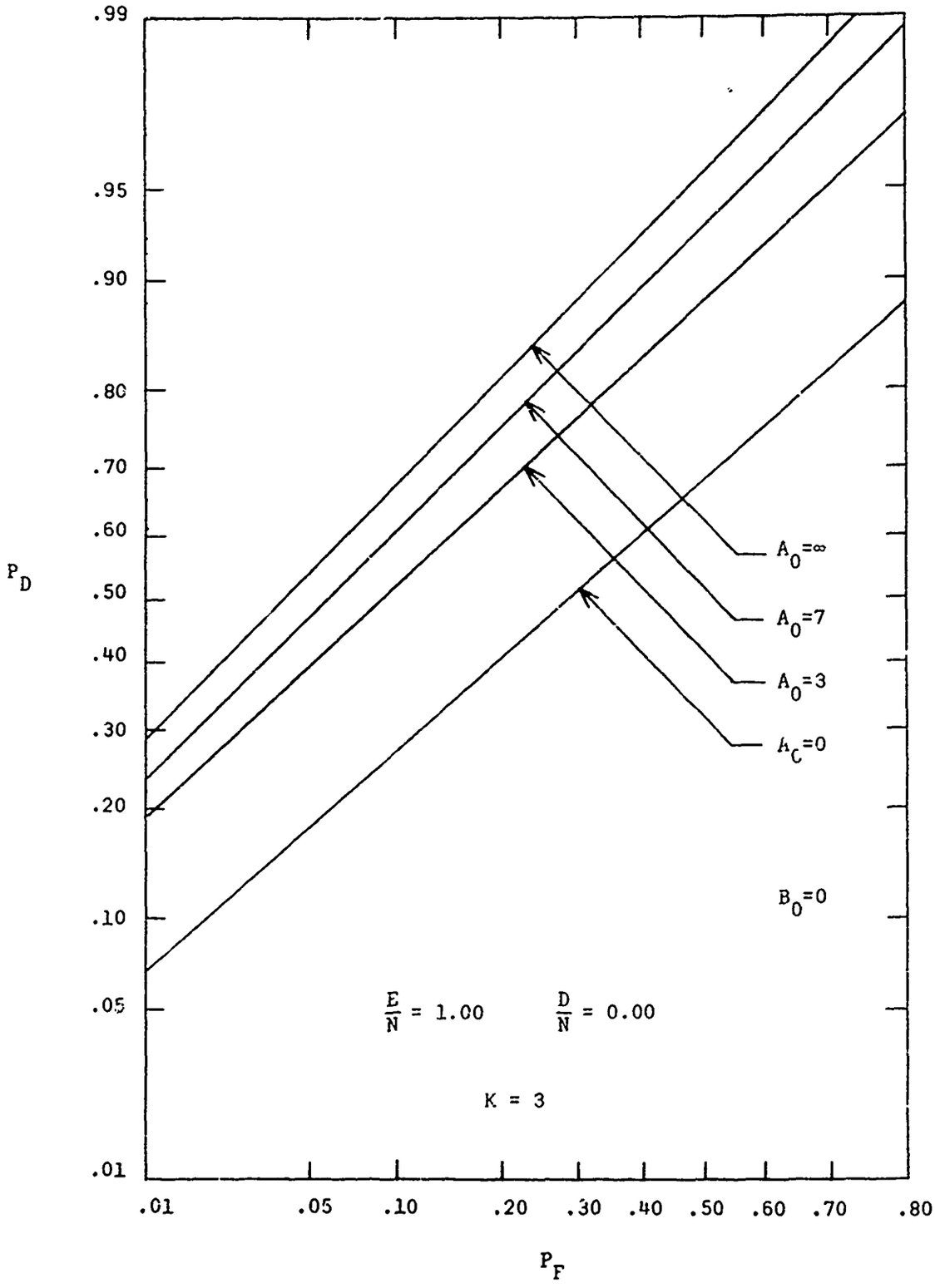


Figure 8.10. Performance of the Suboptimal SKED in NKD Array Processor.

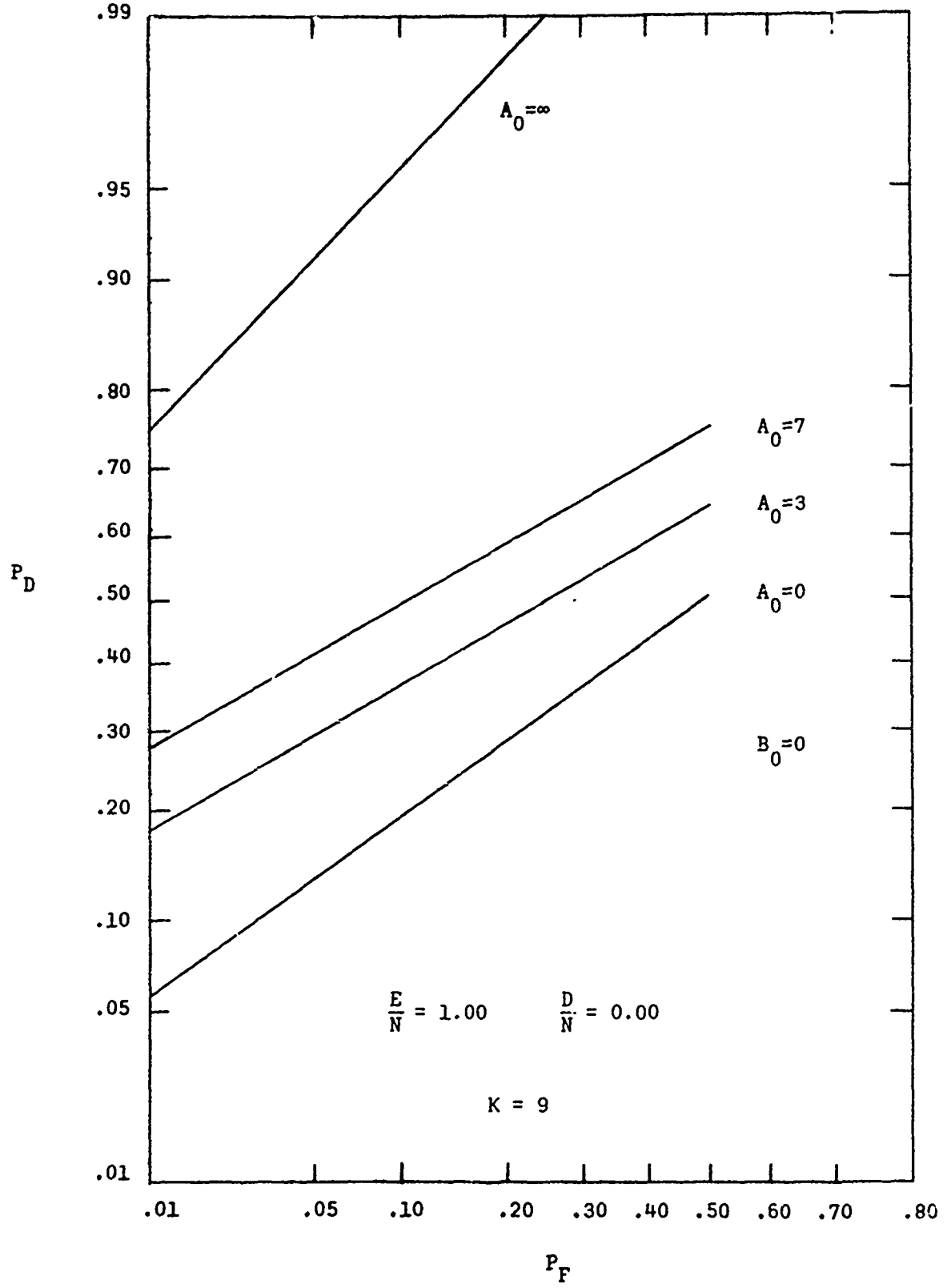


Figure 8.11. Performance of the Suboptimal SKED in NKD Array Processor.

An interesting comparison can be made between the optimal performance just discussed and that for the similar GUD problem reported in Figures 8.2-8.4. If performance degradation was due to spatial uncertainty only, we would expect both problems to suffer a similar performance loss as the level of uncertainty increased. Note that the GUD processor is clearly less effected by location uncertainty than the SKED processor. Thus, spatial and temporal uncertainty get mixed together. In effect, it appears that the Gaussian signal has such a large amount of uncertainty already inherent within it that the addition of location uncertainty does little to degrade detection performance.

The remaining figures in this section are for a directional noise source of unit spectral height. In Figures 8.12-8.16, its location is fixed such that $N\omega_0\tau_n = \pi/2$. Figures 8.12-8.14 correspond to Figures 8.7-8.9. Note that the addition of a directional noise source causes additional performance degradation for all levels of uncertainty. This effect is less noticeable in the nine than the three element array processor.

Figures 8.15 and 8.16 evaluate performance for a particular suboptimal processor which has been placed in the same environment of uncertainty as discussed above. This suboptimal processor has the estimate and plug structure as illustrated in Figure 3.4. Since the uncertain parameter of signal source location exists under H_1 only, $\hat{\underline{\theta}}_1$ is the sole estimate required. In this case, the estimate is not data dependent and is fixed at $\hat{\underline{\theta}}_1 = (\omega_0\hat{\tau}_s) = 0$. The parameter conditional expression into which $\hat{\underline{\theta}}_1$ is plugged is given by

$$\Lambda(\underline{z}|\underline{\theta}_1) = \prod_{n=0}^{N} \Lambda(\underline{z}(n)|\omega_0\tau_s) \tag{8.6}$$

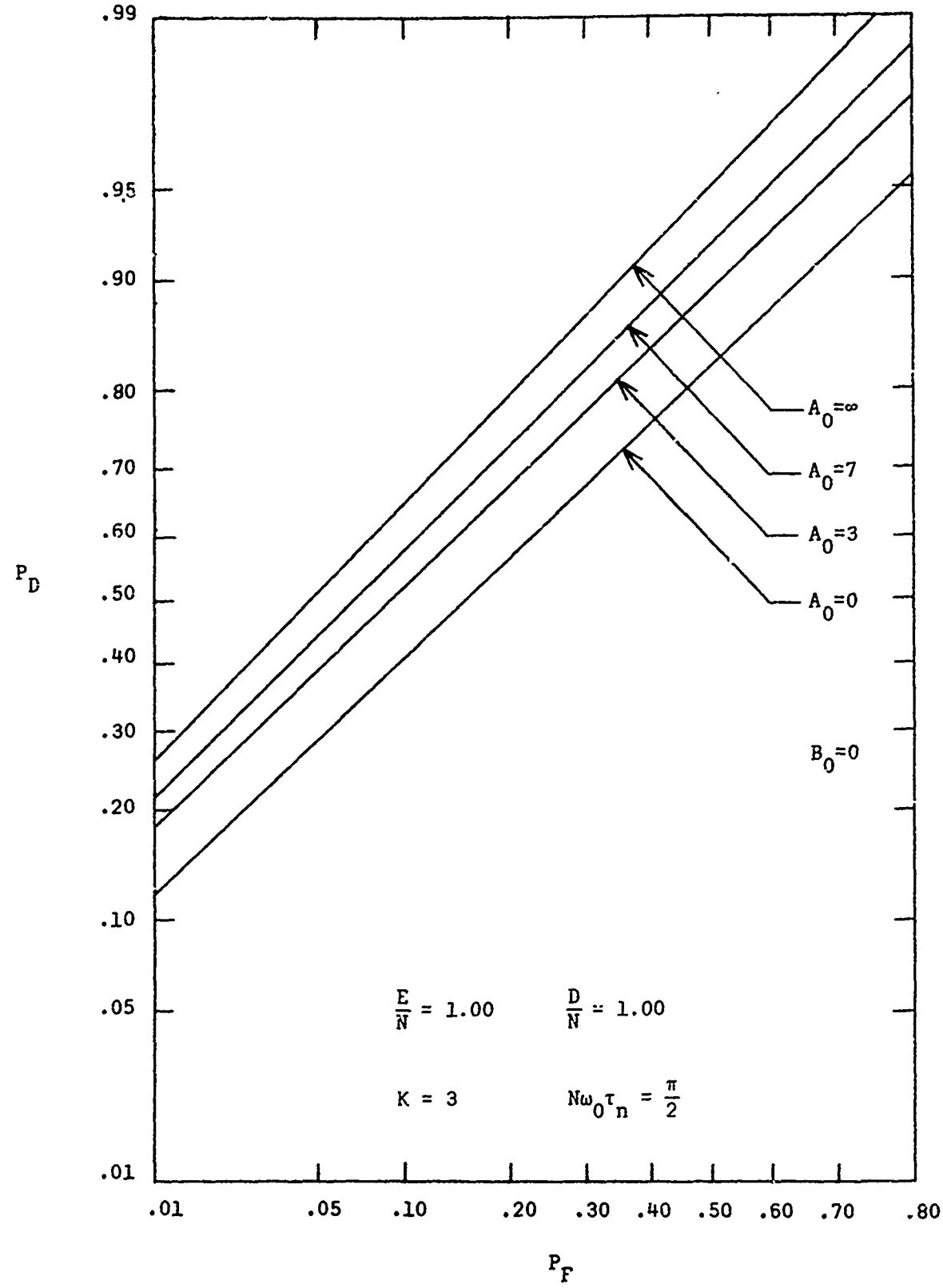


Figure 8.12. Performance of the Optimal SKED in NKD Array Processor.

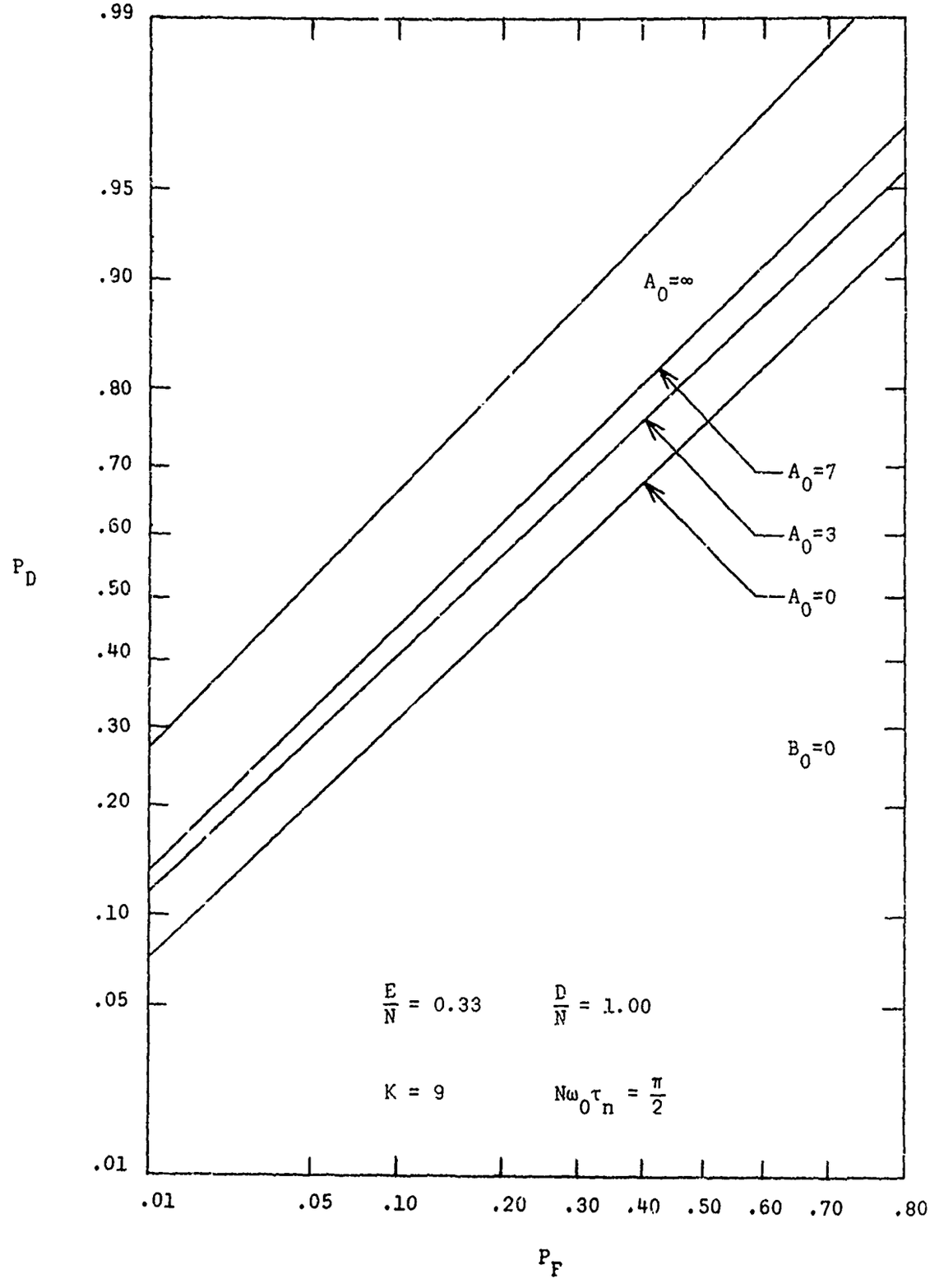


Figure 8.13. Performance of the Optimal SKED in NKD Array Processor.

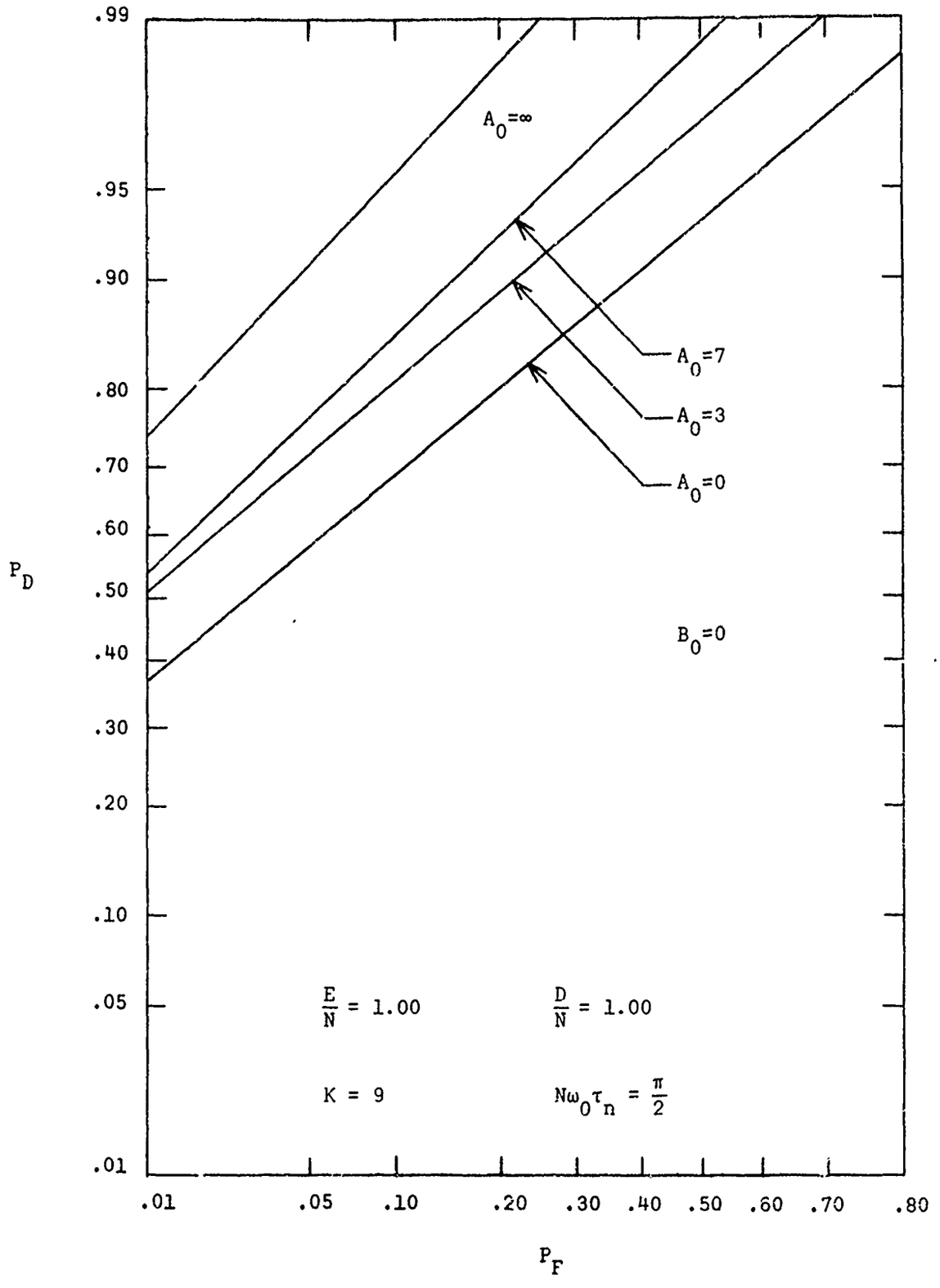


Figure 8.14. Performance of the Optimal SKED in NKD Array Processor.

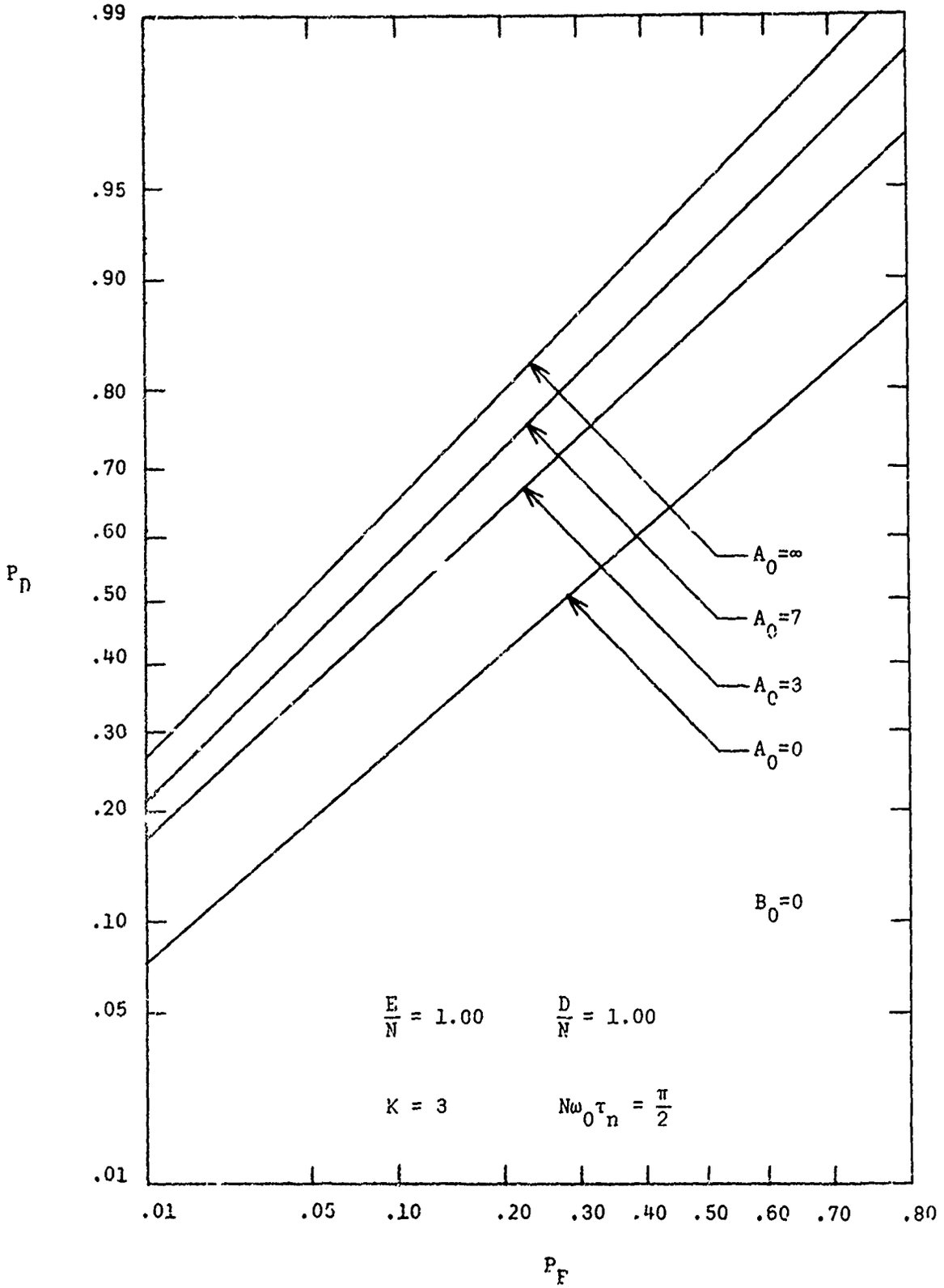


Figure 8.15. Performance of the Suboptimal SKED in Array Processor.

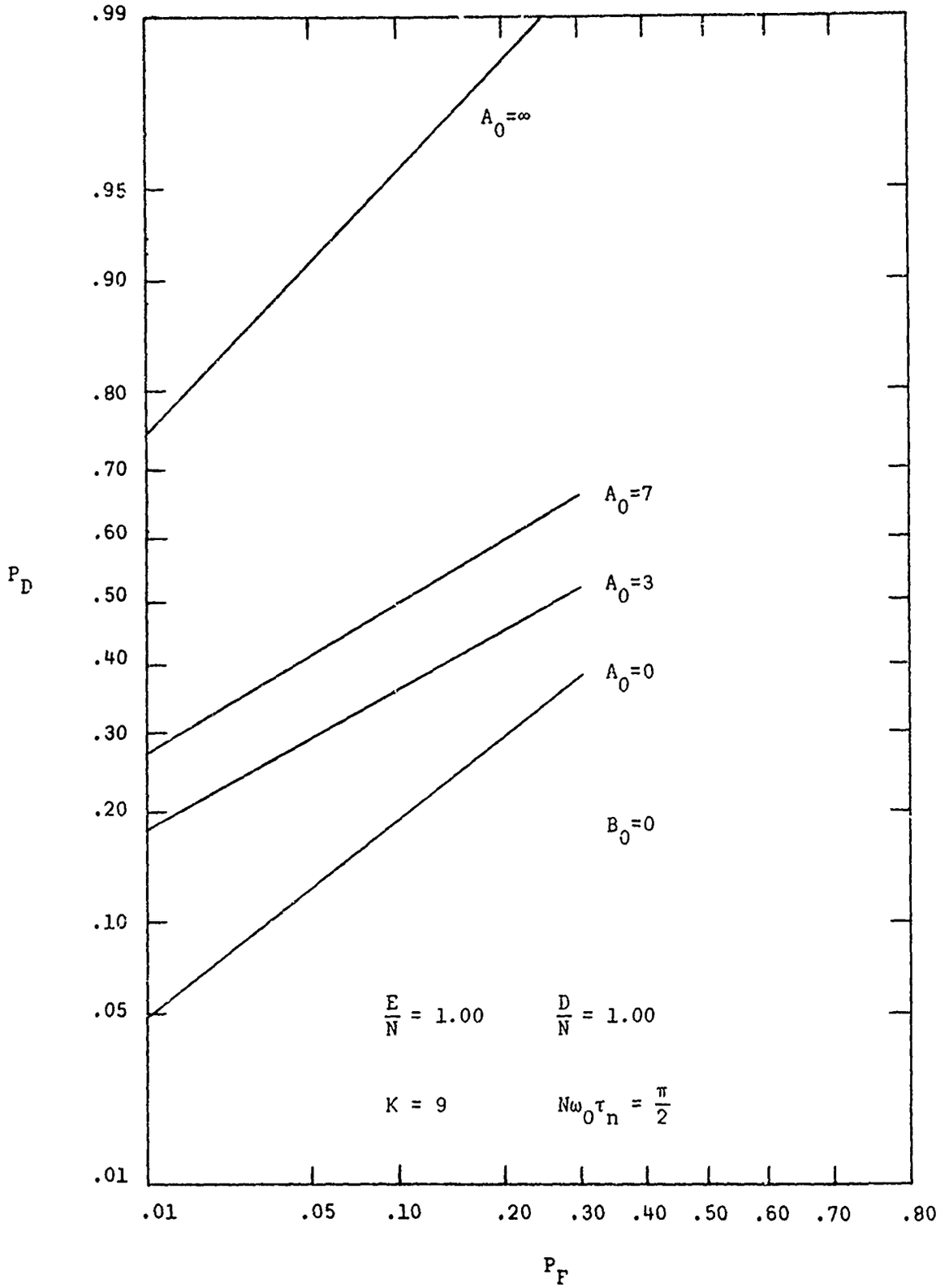


Figure 8.16. Performance of the Suboptimal SKED in NKD Array Processor.

where $\Lambda(\underline{z}(n)|\omega_0\tau_s)$ can be found in Table 6.1 B. The resulting processor is a realization of that derived by Mermoz (Mermoz, 1964; Horton, 1969) with a front end consisting of a beamformer looking broadside to the array and a nullformer pointing towards the directional noise source's location. Note that little difference is noticed between these results and those reported in Figures 8.10 and 8.11. The addition of a noise source of known location has a small performance degradation effect in comparison to the large loss introduced by signal source location uncertainty.

The remaining two figures illustrate the effect of varying the location of the directional noise source for the optimal three element array processor. Figures 8.17, 8.12, and 8.18 correspond to noise source locations such that $N\omega_0\tau_n = 3\pi/4$, $\pi/2$, and $\pi/4$, respectively. In this sequence, the noise source's location is progressively approaching broadside to the array which also is the mean value of the a priori distribution on signal source location. Note that the $A_0 = 0$ curve remains constant while all the remaining curves move closer to it as the difference between the noise source's location and the mean value of the signal source's location becomes smaller.

SKE in NUD

Performance curves are given in Figures 8.19-8.23 for the processor which is to decide presence of absence of a signal known exactly in Gaussian noise consisting of a component which is independent from sensor to sensor and an additive component arising from a source of uncertain direction. Thus, unlike the previous two sections, it is not the location of the signal source which is uncertain. In addition, the resulting processor is more complex than the previous two since averaging over the prior knowledge must take place in both the numerator and denominator of the likelihood ratio as shown below

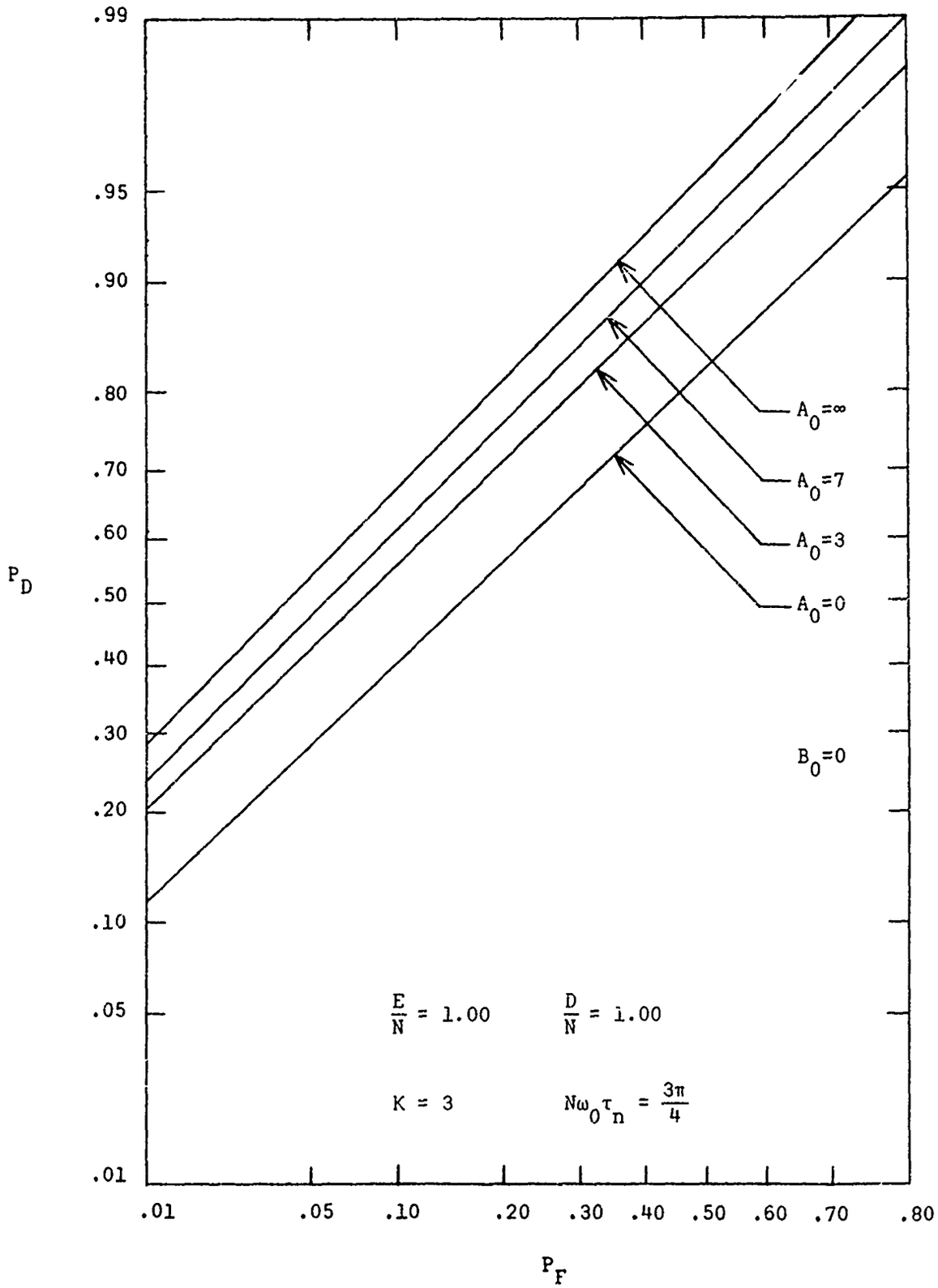


Figure 8.17. Performance of the Optimal SKED in NKD Array Processor.

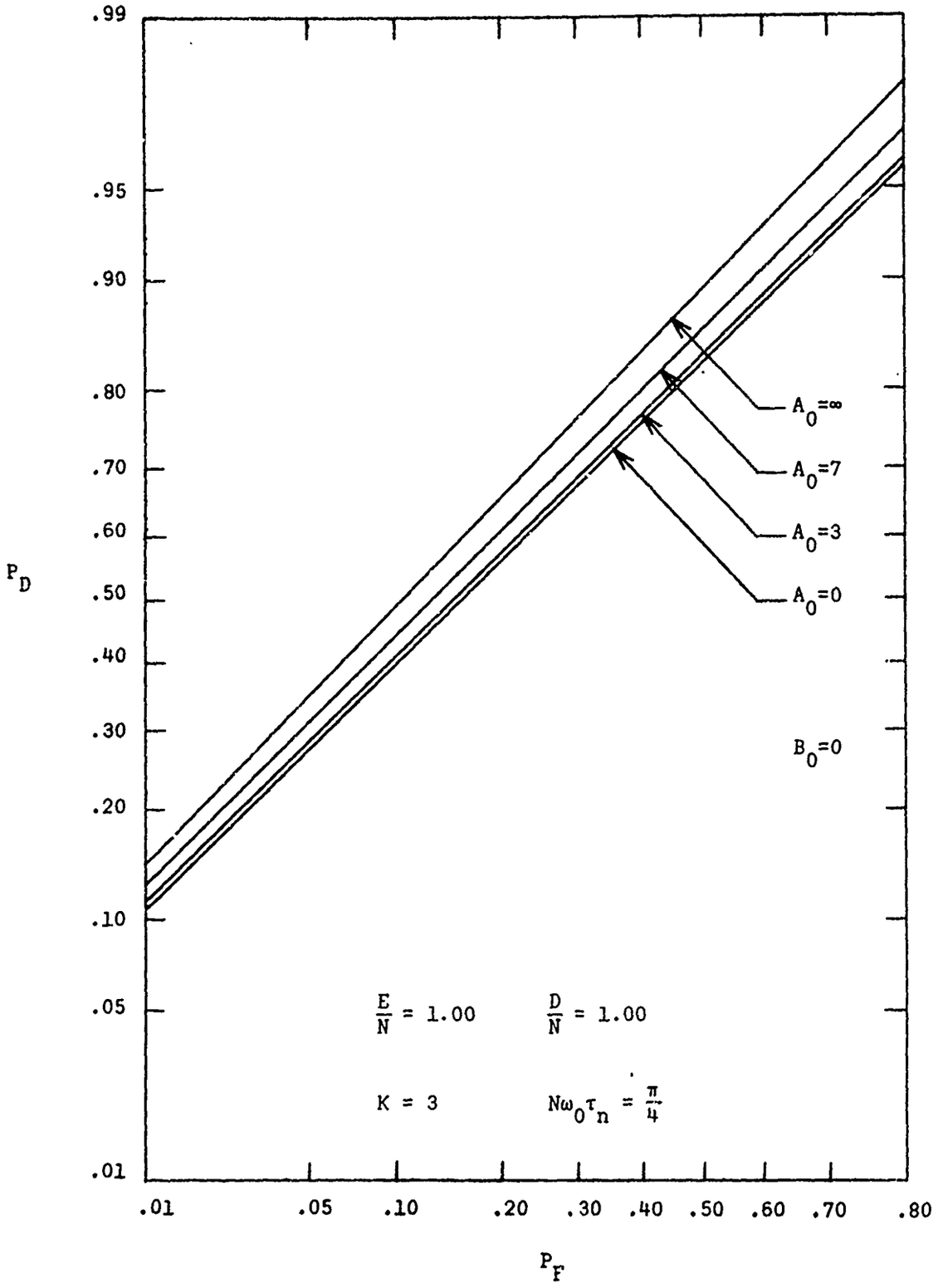


Figure 8.18. Performance of the Optimal SKED in NKD Array Processor.

$$\Lambda(\underline{z}) = \frac{\int_{\omega_0\tau_n} \prod_{n=0}^{N} p(\underline{z}(n)|\omega_0\tau_n, H_1) p(\omega_0\tau_n) d\omega_0\tau_n}{\int_{\omega_0\tau_n} \prod_{n=0}^{N} p(\underline{z}(n)|\omega_0\tau_n, H_0) p(\omega_0\tau_n) d\omega_0\tau_n} \tag{8.7}$$

where $p(\underline{z}(n)|\omega_0\tau_n, H_1)$ and $p(\underline{z}(n)|\omega_0\tau_n, H_0)$ can be found in Table 6.1 C. A single frequency signal is assumed at $N\omega_0/2\pi$ Hz. Its energy over the observation interval is given by $E = 2\, b_0(N)^* b_0(N)$. Note that the figures in this section correspond directly to Figure 8.12-8.16 for the SKED in NKD problem with the locations of the signal and noise source reversed (i.e., $N\omega_0\tau_s = \pi/2$). Thus, it will be possible to compare the relative effects of signal versus noise source location uncertainty.

An interesting contrast with previous results is seen in Figures 8.19-8.21. Note that increasing location uncertainty does not lead to a greater degradation in performance the larger the array size. This is particularly evident in Figures 8.19 and 8.20 where the total signal energy processed has been held constant. An actual decrease in performance loss with increasing array size is seen.

A second contrast with previous results is seen in Figures 8.22 and 8.23. They evaluate performance for a particular suboptimal processor which has been placed in the same environment of uncertainty as discussed above. This processor has the estimate and plug structure as illustrated in Figure 3.4. Since the uncertain parameter of noise source location exists under H_1 and H_0, both the $\hat{\underline{\theta}}_1$ and $\hat{\underline{\theta}}_0$ estimates are required. In this case, the estimates are not data dependent and are fixed at $\hat{\underline{\theta}}_1 = \hat{\underline{\theta}}_0 = (\omega_0\tau_n) = 0$. The parameter conditional expression into which $\hat{\theta}_1$ and $\hat{\underline{\theta}}_0$ are plugged is given by

$$\Lambda(\underline{z}|\underline{\theta}_1,\underline{\theta}_0) = \frac{\prod_{n=0}^{N} p(\underline{z}(n)|\omega_0\tau_n,H_1)}{\prod_{n=0}^{N} p(\underline{z}(n)|\omega_0\tau_n,H_0)} \tag{8.8}$$

where $p(\underline{z}(n)|\omega_0\tau_n,H_1)$ and $p(\underline{z}(n)|\omega_0\tau_n,H_0)$ can be found in Table 6.1 C. Similar to the previous section, the resulting processor is a realization of that derived by Mermoz (Mermoz, 1964; Horton, 1969) with a front end consisting of both a beamformer and nullformer. The beamformer is looking towards the signal source's location and the nullformer is pointing broadside to the array. Note that while significant performance degradation is suffered for all levels of uncertainty other than precise knowledge, the loss is nowhere near as severe as in Figures 8.15 and 8.16 where it is the signal source's location which is uncertain.

Performance Summary

The results of this chapter facilitate an understanding of the effect of both signal and noise source location uncertainty on array processor performance. Several observations have been made and they are summarized below.

The detectabilities of a known form and Gaussian signal of uncertain location were compared. While both exhibited a degradation in performance as the level of uncertainty increased, it was noted that the optimal array processor for a Gaussian signal suffered significantly less in this respect than that for a known form signal.

The remaining comparisons between optimal array processors were for problems involving both a known form directional signal source and a directional noise source. A question that arises immediately is which of the two sources has a greater effect on detectability when its location is uncertain. For the E/N and D/N ratios investigated, it was found that

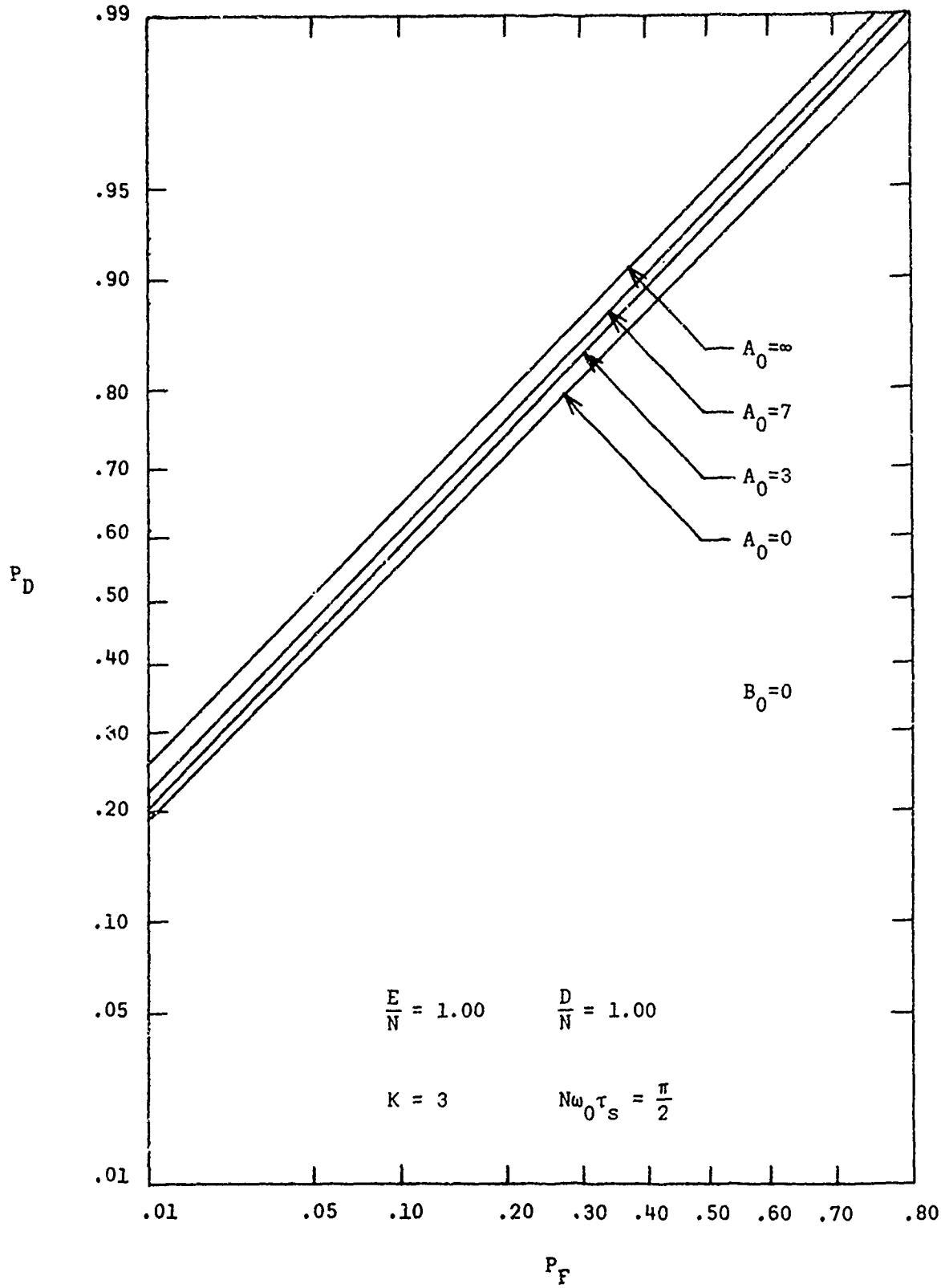


Figure 8.19. Performance of the Optimal SKE in NUD Array Processor.

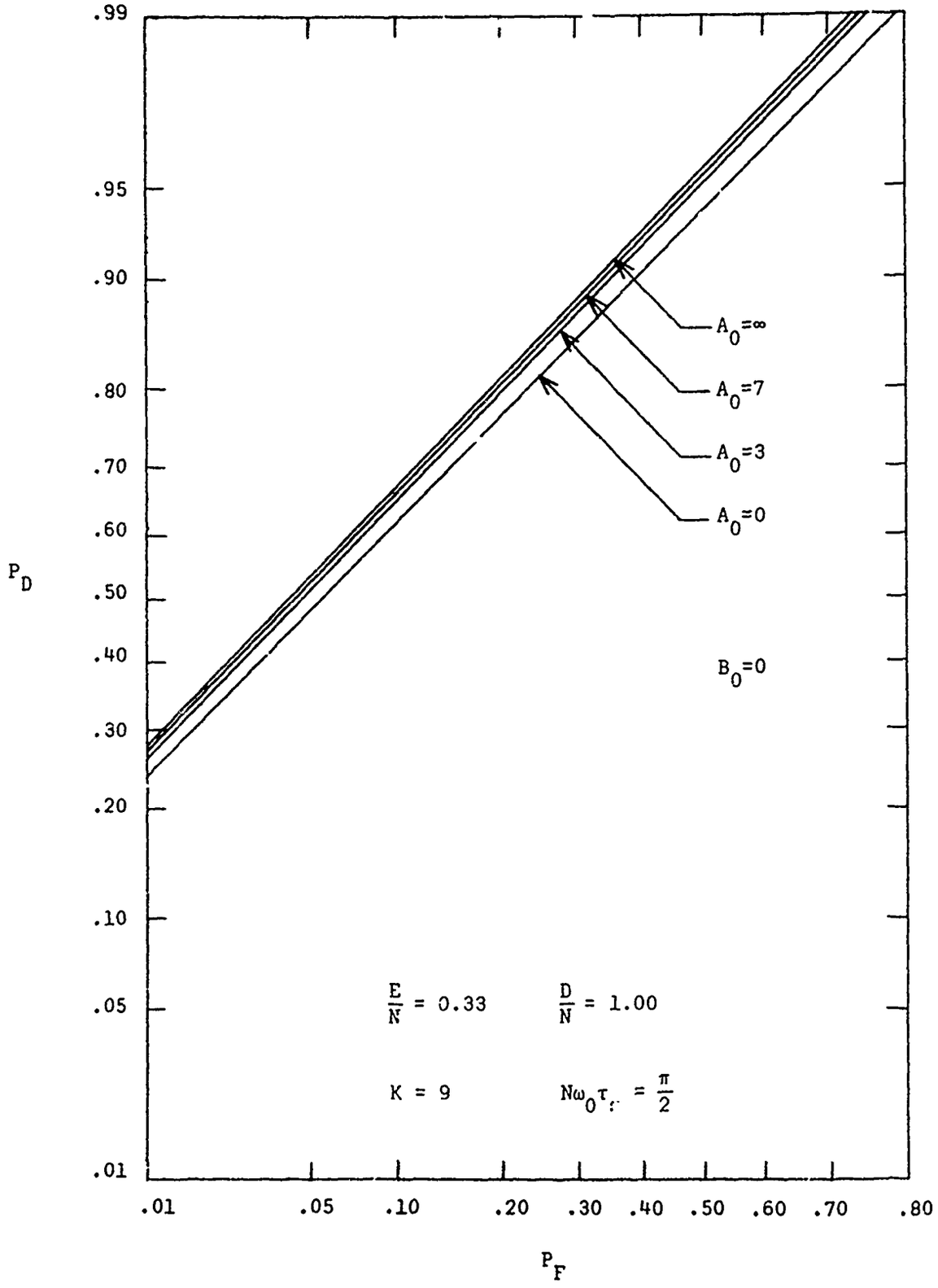


Figure 8.20. Performance of the Optimal SKE in NUD Array Processor.

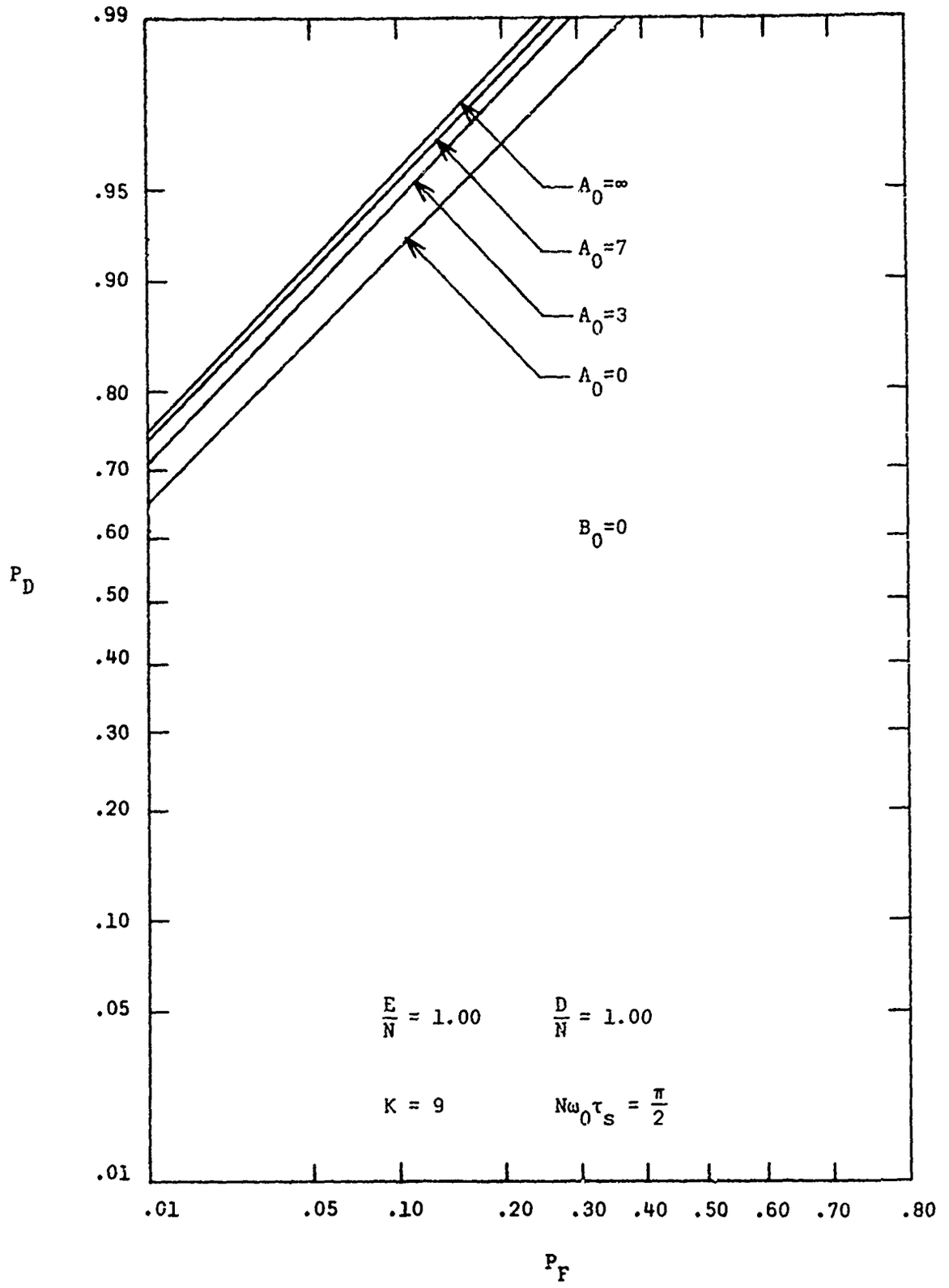


Figure 8.21. Performance of the Optimal SKE in NUD Array Processor.

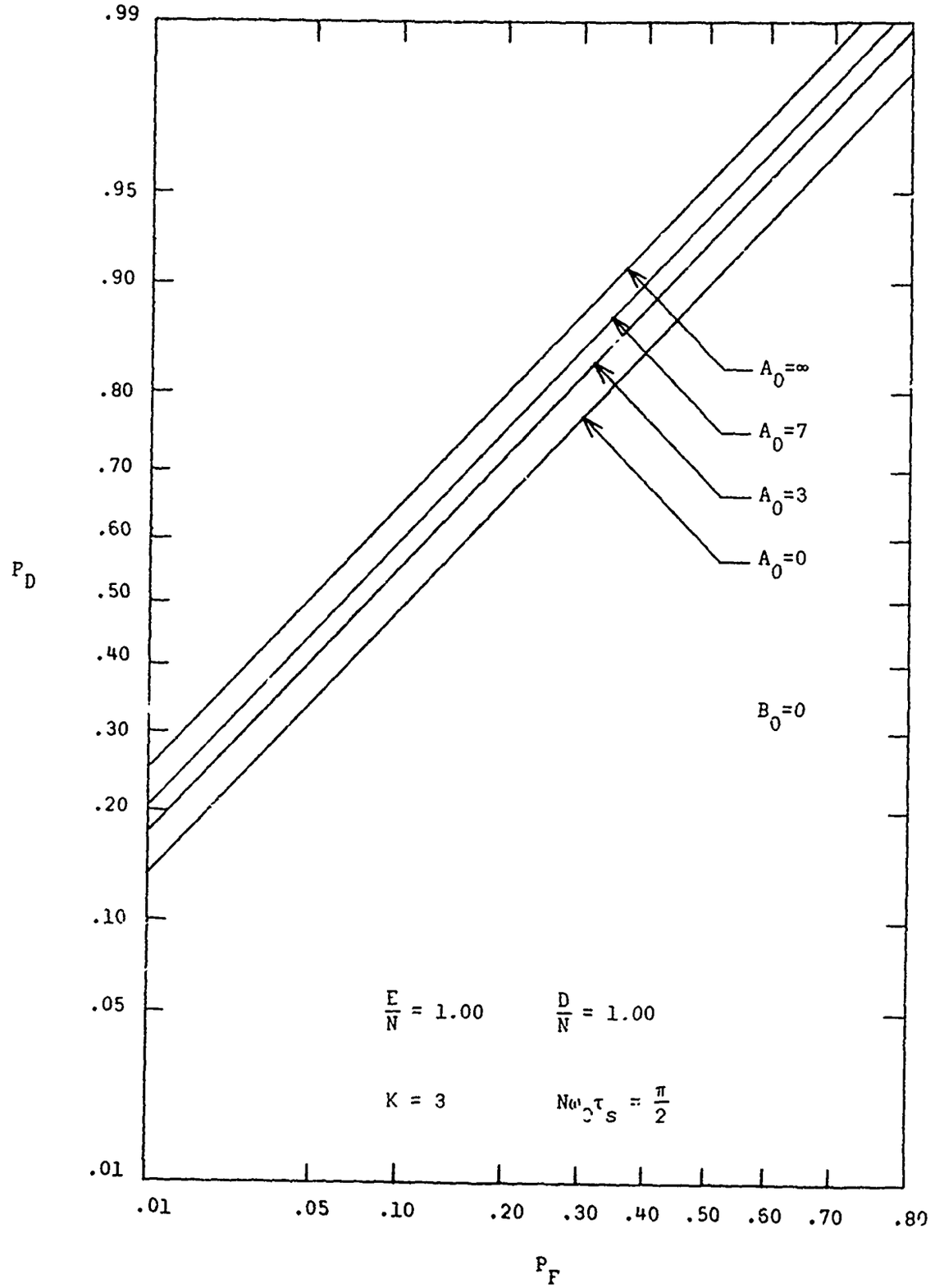


Figure 8.22. Performance of the Suboptimal SKE in NUD Array Processor.

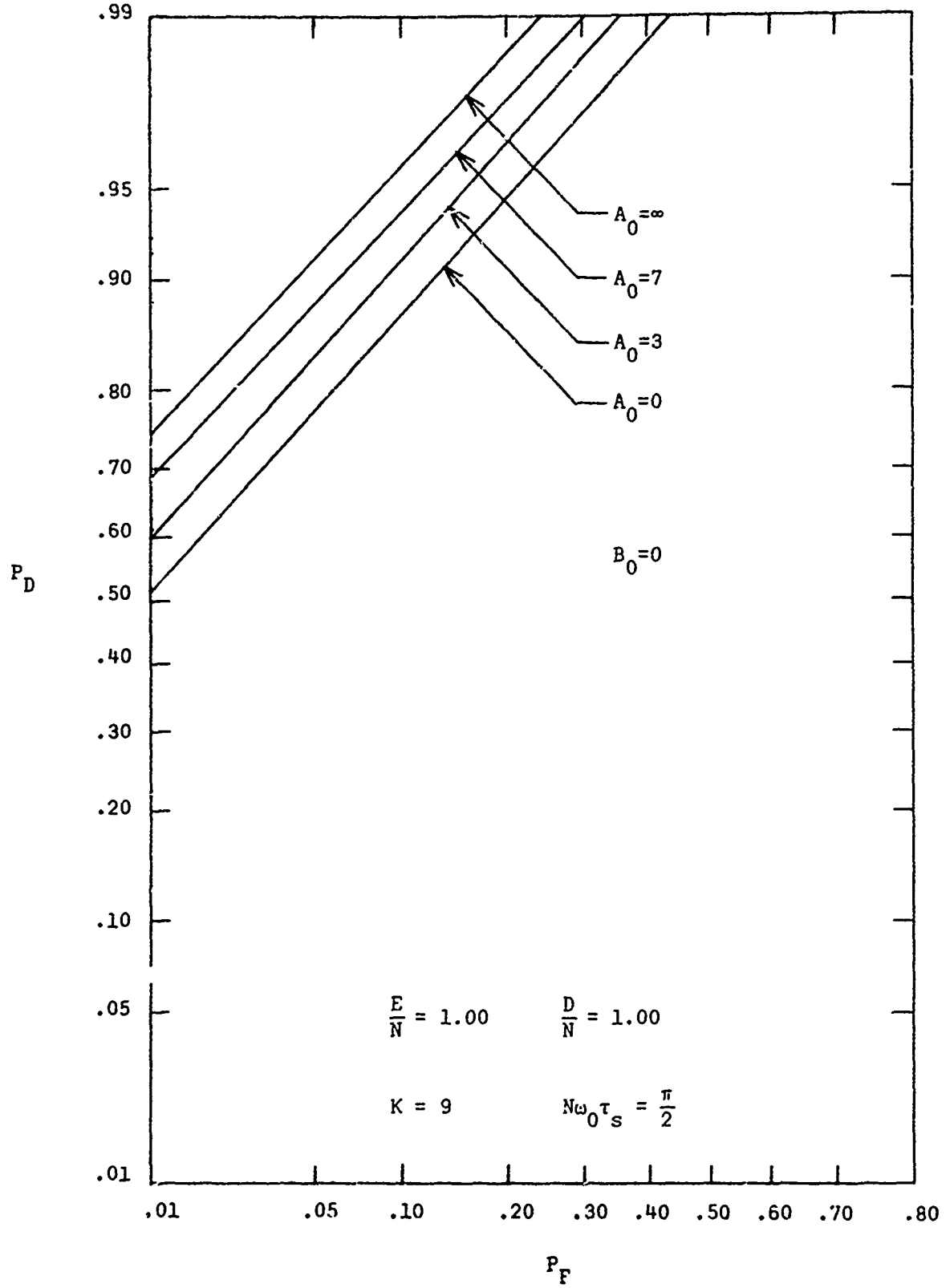


Figure 8.23. Performance of the Suboptimal SKE in NUD Array Processor.

uncertainty in the location of a noise source has a less serious effect on performance degradation than a corresponding amount of uncertainty on the location of a known form signal source.

Since performance results for two array sizes were investigated, the effect of increasing the array size while keeping the level of uncertainty constant could be studied. When there is signal source location uncertainty, increasing the array size leads to a degradation in performance. In contrast, when it is the noise source's location which is uncertain, increasing the array size does not lead to a degradation in performance. In fact, a slight increase was noted.

An appealing approach to array processor design when uncertain parameters exist is first to estimate these parameters, then plug them into the parameters known likelihood ratio as if they were known exactly. The performance of a particular class of such suboptimal processors operating in the same environment of uncertainty as their optimal counterparts was investigated. For this class, the estimates of signal and noise source location were not data dependent and fixed at zero. The resulting processors were realizations of those derived by Bryn (Bryn, 1962) and Mermoz (Mermoz, 1964; Horton, 1969) each with a front end consisting of a beamformer and (when appropriate) a nullformer. In all cases, the suboptimal processor suffered greater performance degradation for a given level of uncertainty other than precise knowledge than its optimal counterpart. Especially as the array size grows larger, the results point out the necessity of properly incorporating *a priori* knowledge into array processor design.

Chapter IX

ADAPTIVE LEARNING OF AN ADDITIVE DIRECTIONAL NOISE COMPONENT'S LOCATION

As mentioned at the end of Chapter III, an appealing approach to an array detection problem where uncertain parameters exist is to estimate these parameters and then plug them into the conditional likelihood ratio as if they were known exactly. When "good" estimators are used in the structure illustrated in Figure 3.4, the processor is typically referred to as being adaptive. Once again, however, it is not clear that piecing together locally optimal techniques will yield global optimality when the overall goal is good detection performance. The intent of this chapter is to reiterate that when implemented sequentially, the optimal array processor exhibits learning or adaptive features naturally. Specific cases of the SKE in NUD processor will be used as illustrations.

Adaptive Optimal Array Processors

In the formation of the likelihood ratio, the marginal distribution of the observables conditional to each hypothesis is needed. Assume the total observation period has been broken into a sequence of incremental observation periods. Adopting the notation of Chapter V, our observables are now the L vectors $(\underline{z}^1,\ldots,\underline{z}^L)$. Suspending the conditioning to H_1 and H_0

$$p(\underline{z}^1,\ldots,\underline{z}^L) = \prod_{i=1}^{L} p(\underline{z}^i|\underline{z}^{i-1},\ldots,\underline{z}^1). \tag{9.1}$$

Assuming parameter conditional independence of the $\underline{z}^i$ (see Chapter V)

$$p(\underline{z}^i|\underline{z}^{i-1},\ldots,\underline{z}^1) = \int_\theta p(\underline{z}^i|\underline{z}^{i-1},\ldots,\underline{z}^1,\underline{\theta})p(\underline{\theta}|\underline{z}^{i-1},\ldots,\underline{z}^1)d\underline{\theta}$$

$$= \int_\theta p(\underline{z}^i|\underline{\theta})p(\underline{\theta}|\underline{z}^{i-1},\ldots,\underline{z}^1)d\underline{\theta} \qquad (9.2)$$

where $p(\underline{\theta}|\underline{z}^{i-1},\ldots,\underline{z}^1)$ is an updated version of the a priori probability density function of $\underline{\theta}$

$$p(\underline{\theta}|\underline{z}^{i-1},\ldots,\underline{z}^1) = \frac{p(\underline{z}^{i-1}|\underline{\theta})p(\underline{\theta}|\underline{z}^{i-2},\ldots,\underline{z}^1)}{p(\underline{z}^{i-1}|\underline{z}^{i-2},\ldots,\underline{z}^1)} . \qquad (9.3)$$

The expressions (9.1), (9.2) and (9.3) are the sequential design equations used to form the marginal distributions required to calculate the likelihood ratio

$$\Lambda(\underline{z}^1,\ldots,\underline{z}^L) = \frac{p(\underline{z}^1,\ldots,\underline{z}^L|H_1)}{p(\underline{z}^1,\ldots,\underline{z}^L|H_0)} . \qquad (9.4)$$

Figure 5.1 illustrates the corresponding array processor structure. The adaptive feature arises out of the sequential Bayesian updating of the a priori knowledge of the uncertain parameter vector. In general, the numerator and denominator equations in (9.4) must remain separated in the updating sequence.

An Example: SKE in NUD

In this section, computer simulations of the SKE in NUD processor for several cases will be used as illustrations of the natural adaptive feature of an optimal array processor when implemented sequentially. The uncertain parameter in this case is the Gaussian noise source's location reflected in the phase term $\underline{\theta} = \omega_0\tau_n$. The parameter conditional joint density expressions under H_1 and H_0 required in (9.2) and (9.3) are given by

$$p(\underline{z}^i|\underline{\theta},H_1) = \prod_{n=0}^{N} p(\underline{z}^i(n)|\omega_0\tau_n,H_1) \tag{9.5}$$

$$p(\underline{z}^i|\underline{\theta},H_0) = \prod_{n=0}^{N} p(\underline{z}^i(n)|\omega_0\tau_n,H_0) \tag{9.6}$$

where $p(\underline{z}^i(n)|\omega_0\tau_n,H_1)$ and $p(\underline{z}^i(n)|\omega_0\tau_n,H_0)$ can be found in Table 6.1 C. The superscript is omitted in Table 6.1 C since the expressions are the same for all L incremental observation intervals.

Figures 9.1-9.12 illustrate single computer simulation runs of the processor consisting of 13 iterations each. Since the uncertain parameter exists under both hypotheses, two columns of the sequentially updated a priori knowledge corresponding to (9.3) under H_0 and H_1 are recorded. Each probability density function displays P(WOTN) = $p(\omega_0\tau_n|\underline{z}^i,\ldots,\underline{z}^1)$ versus WOTN = $\omega_0\tau_n$ where ITER = i. The first density in each column (ITER = 0) is the a priori knowledge. A uniform prior of $p(\omega_0\tau_n) = 4/\pi$ for $-\pi/8 \leq \omega_0\tau_n \leq \pi/8$ was assumed in all cases. The "MAX" value indicates the maximum value of the particular density recorded. Between the two columns of densities is a third column which indicates the iteration number ("ITER") and the value of the likelihood ratio ("L") at the completion of that iteration as given by (9.4) utilizing (9.1). The lower left hand corner of each figure records the incremental signal energy processed ("E"), noise power spectra heights ("N" and "D"), and array size ("K"). The noise spectra are assumed white and bandlimited to $8\omega_0/2\pi$ Hz. The signal consists of a single frequency at $8\omega_0/2\pi$ Hz. Its energy over one incremental observation interval is given by $E = 2b_0(8)^* b_0(8)$. The spacing between the array elements is assumed one half wavelength at the frequency of the signal. The lower right hand corner of each figure records the a priori probability density function parameters ("AO" and "BO", see 1)), the true hypothesis in force

during the simulation ("H1" or "H0"), and the true location of the noise source reflected in terms of phase ("WOTN"). In all cases, WOTN = $\omega_0\tau_n = 0$.

The first six figures report single processor runs for a three element array; the last six do the same for a nine element array. The figures are given in pairs. In the first, the true hypothesis in force is H_0; in the second, H_1. All figures have a signal-to-noise ratio of $E/N = 1$. The noise-to-noise ratio (D/N) has the values .01, .03, and .09. Note that even though the noise-to-noise ratios investigated were relatively low, the optimal processor was usually able to learn the noise source's location (under the correct hypothesis). It is interesting that under the incorrect hypothesis, the sequentially updated a priori knowledge often peaks up at the location of the signal source which in terms of phase was $\omega_0\tau_s = \pi/16$. As these figures indicate, the optimal array processor exhibits natural learning or adaptive features when implemented sequentially.

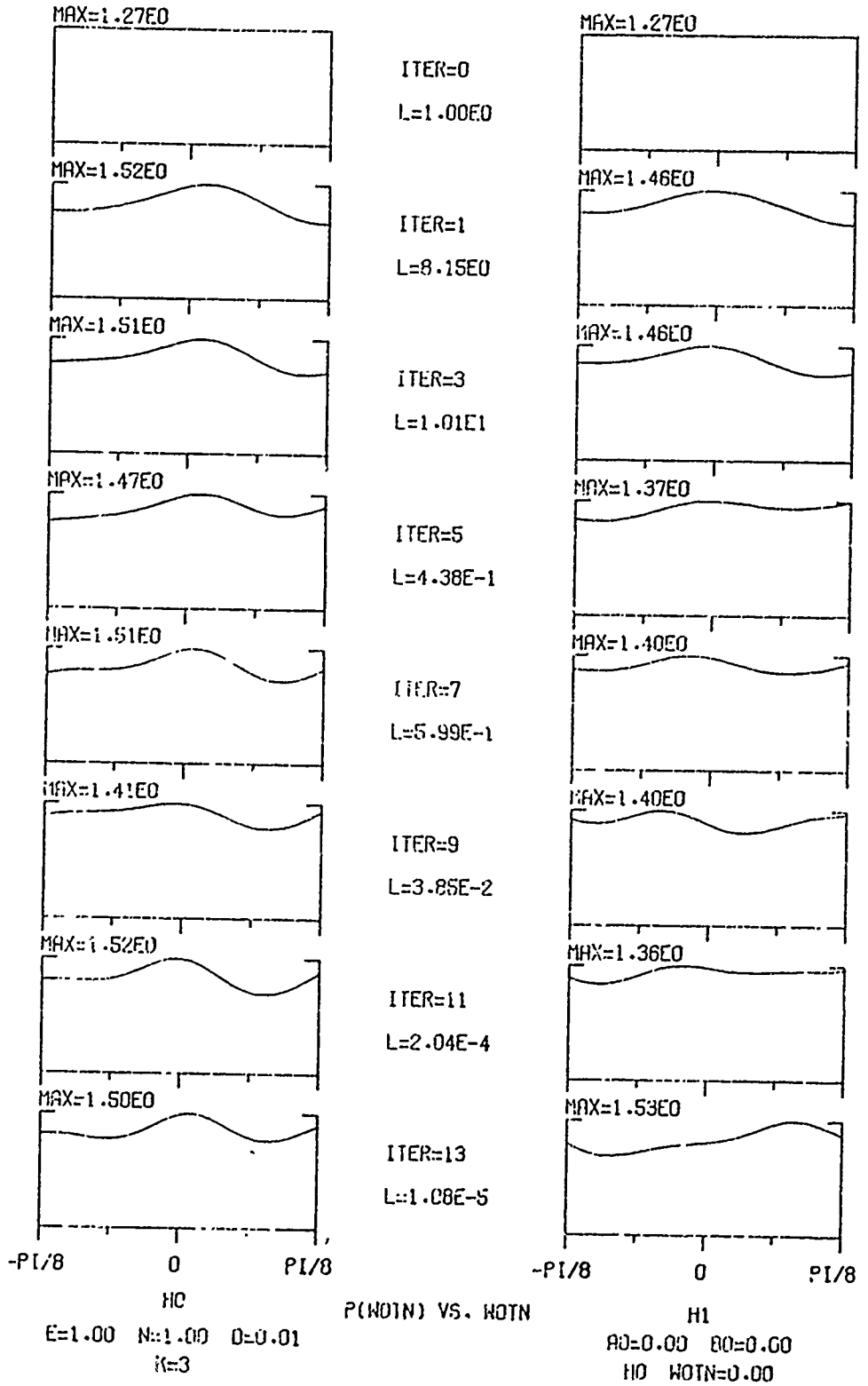


Figure 9.1. Sequential SKE in NUD Simulation.

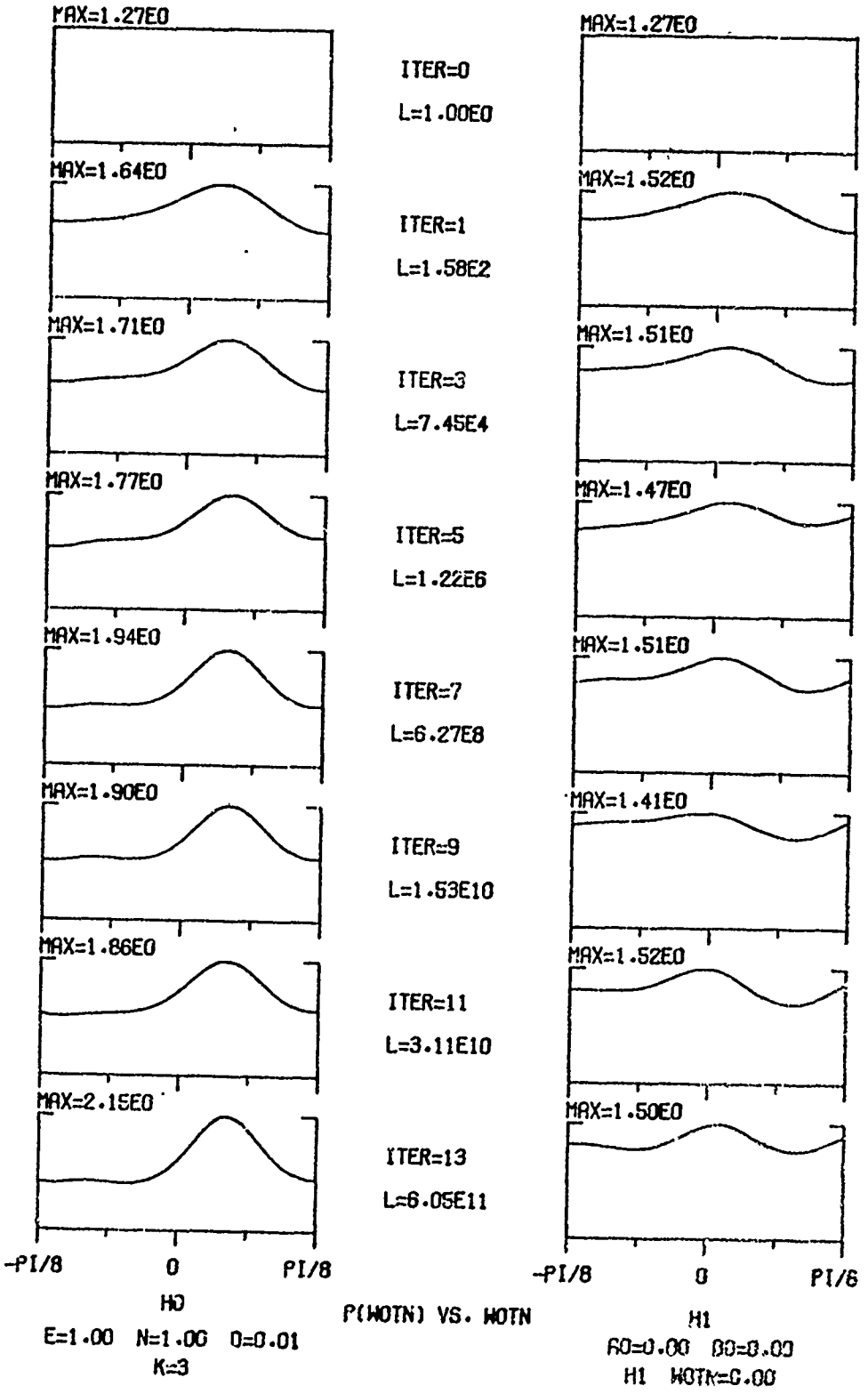


Figure 9.2. Sequential SKE in NUD Simulation.

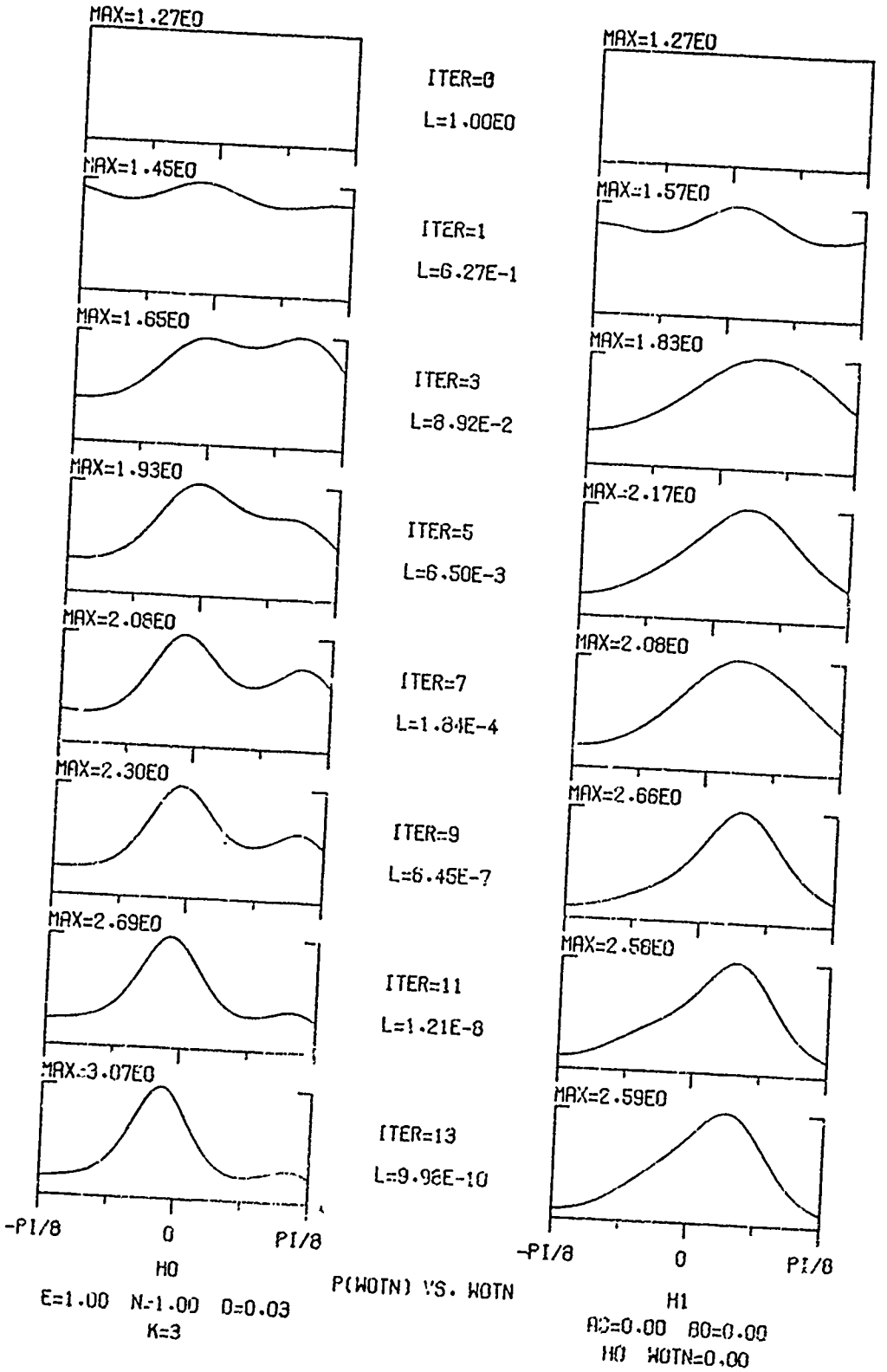


Figure 9.3. Sequential SKE in NUD Simulation.

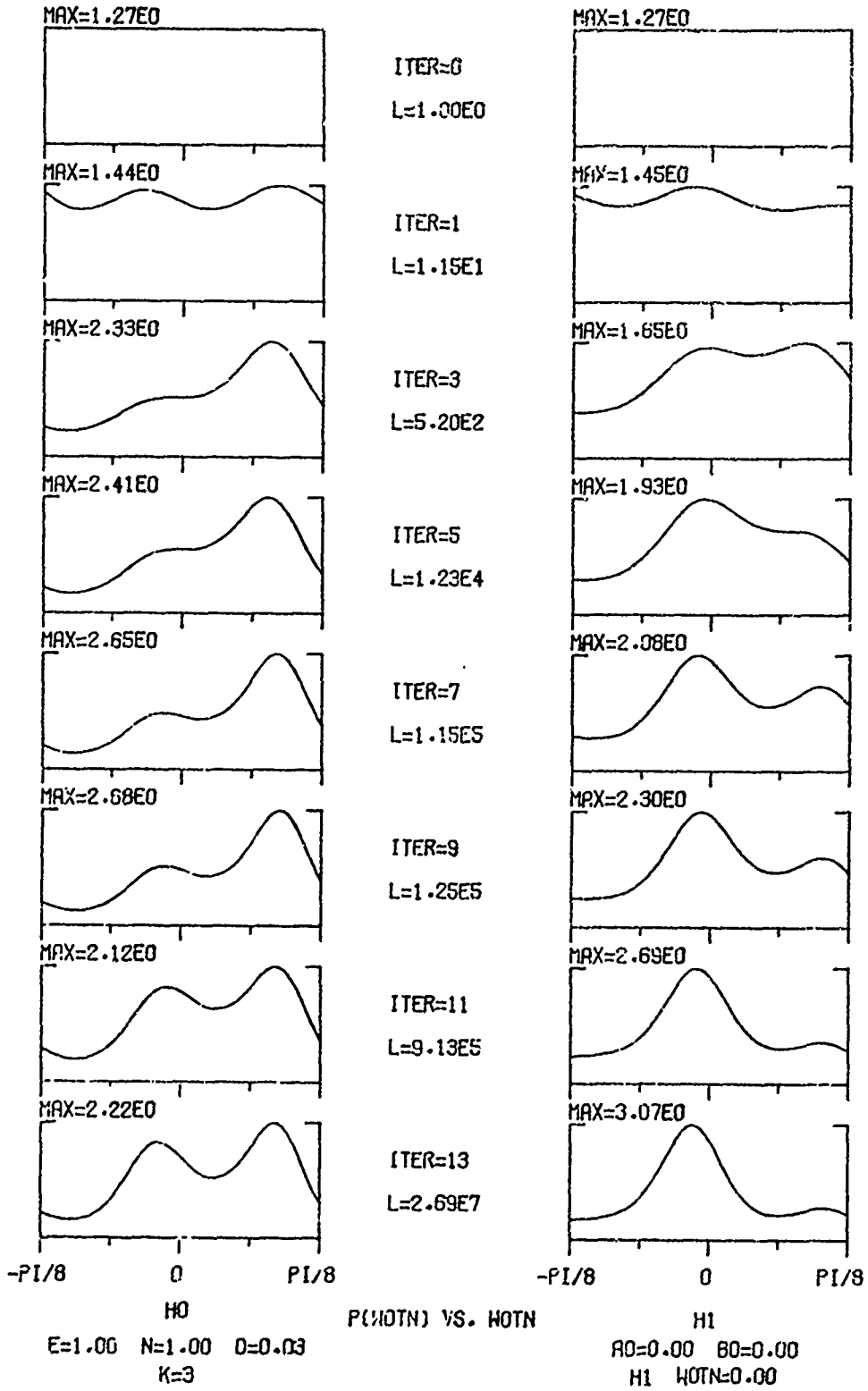


Figure 9.4. Sequential SKE in NUD Simulation.

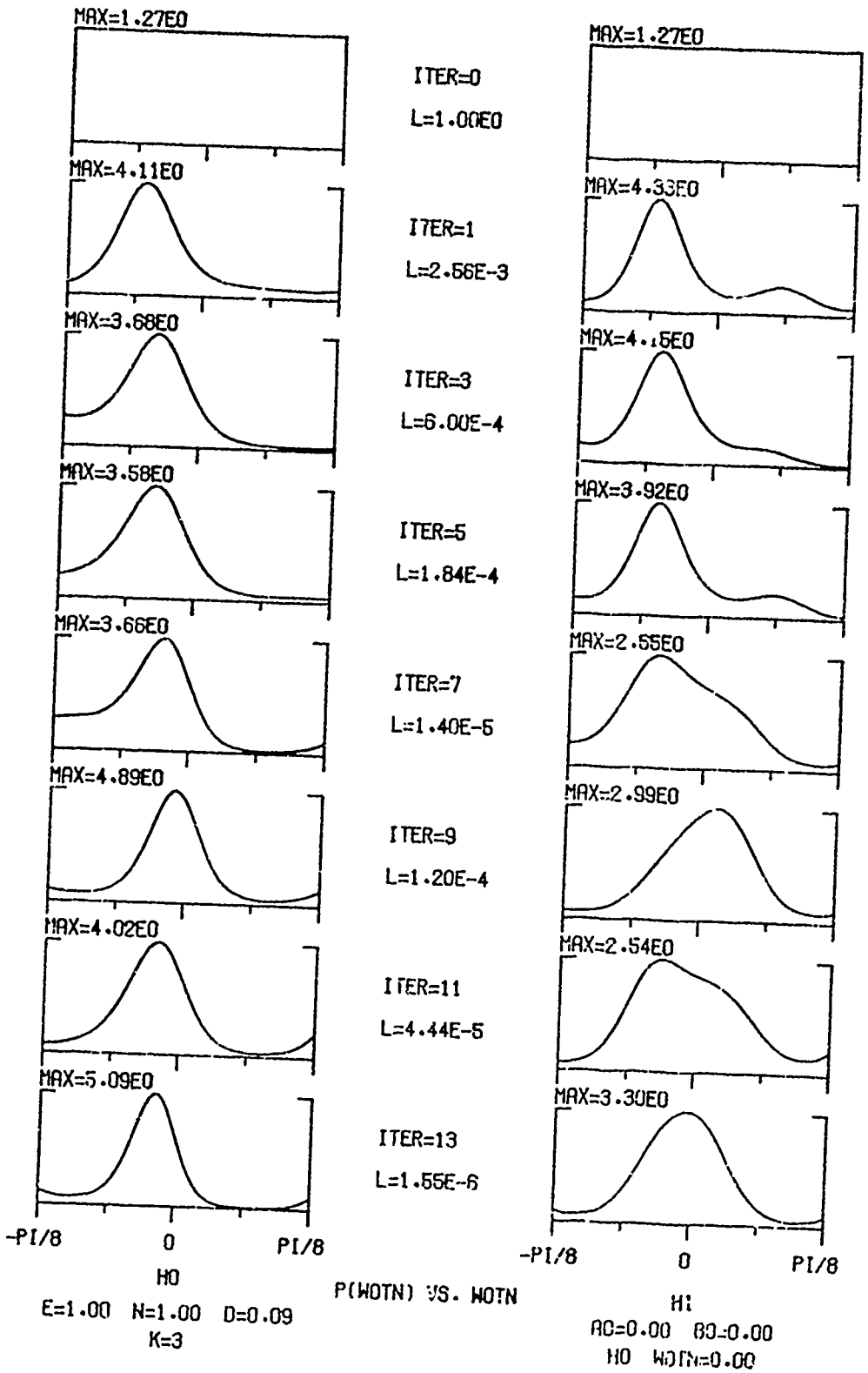


Figure 9.5. Sequential SKE in NUD Simulation.

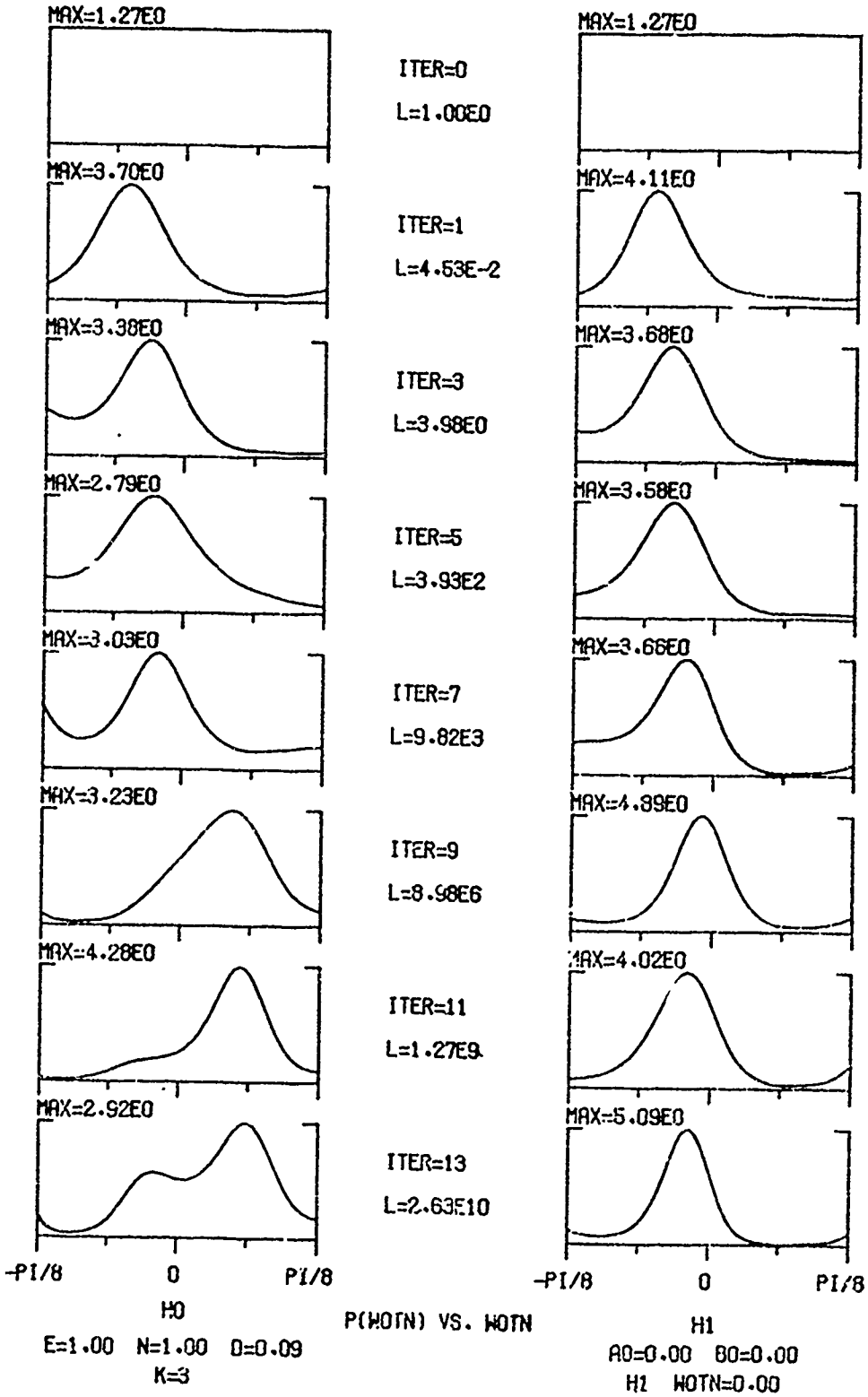


Figure 9.6. Sequential SKE in NUD Simulation.

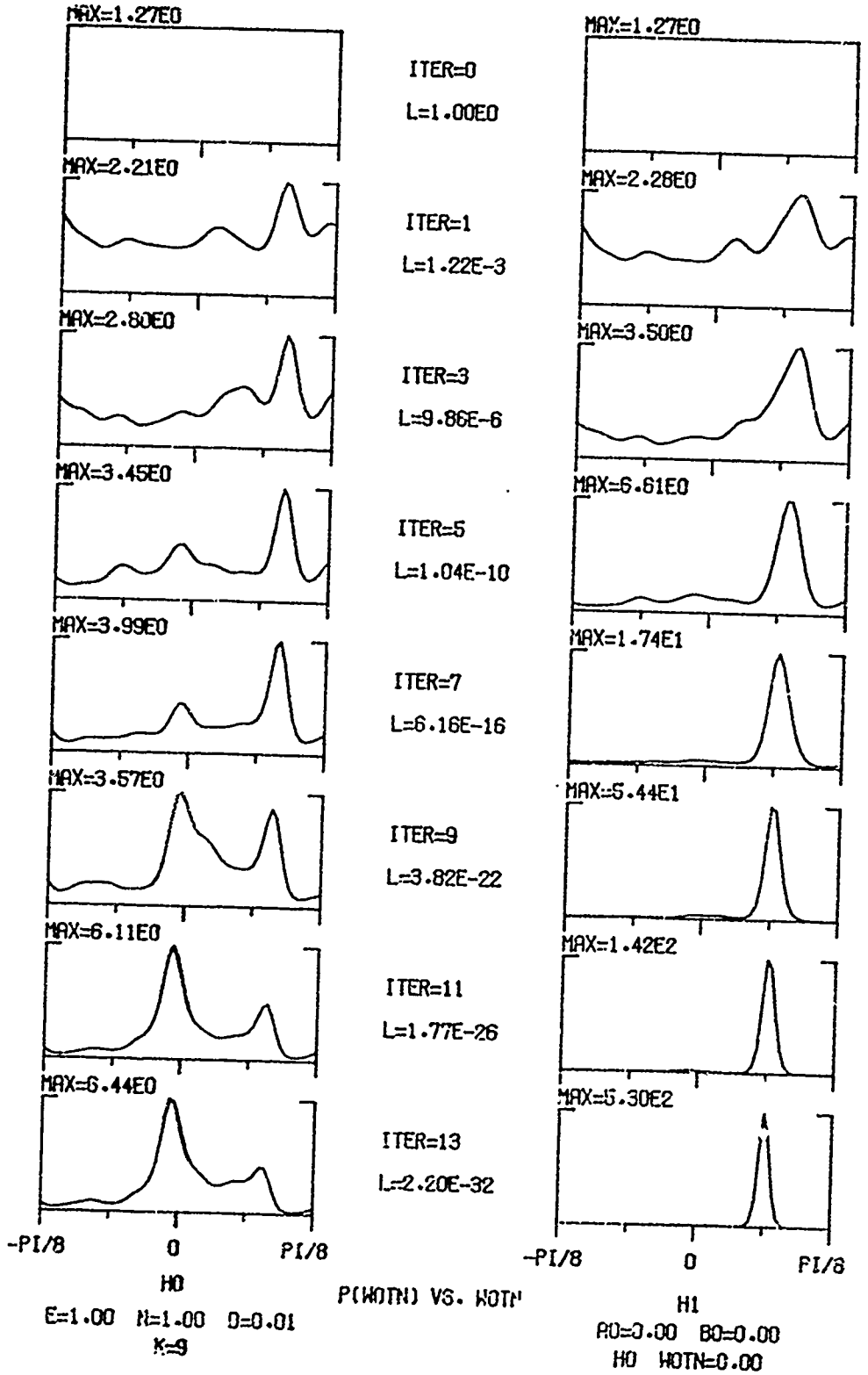


Figure 9.7. Sequential SKE in NUD Simulation.

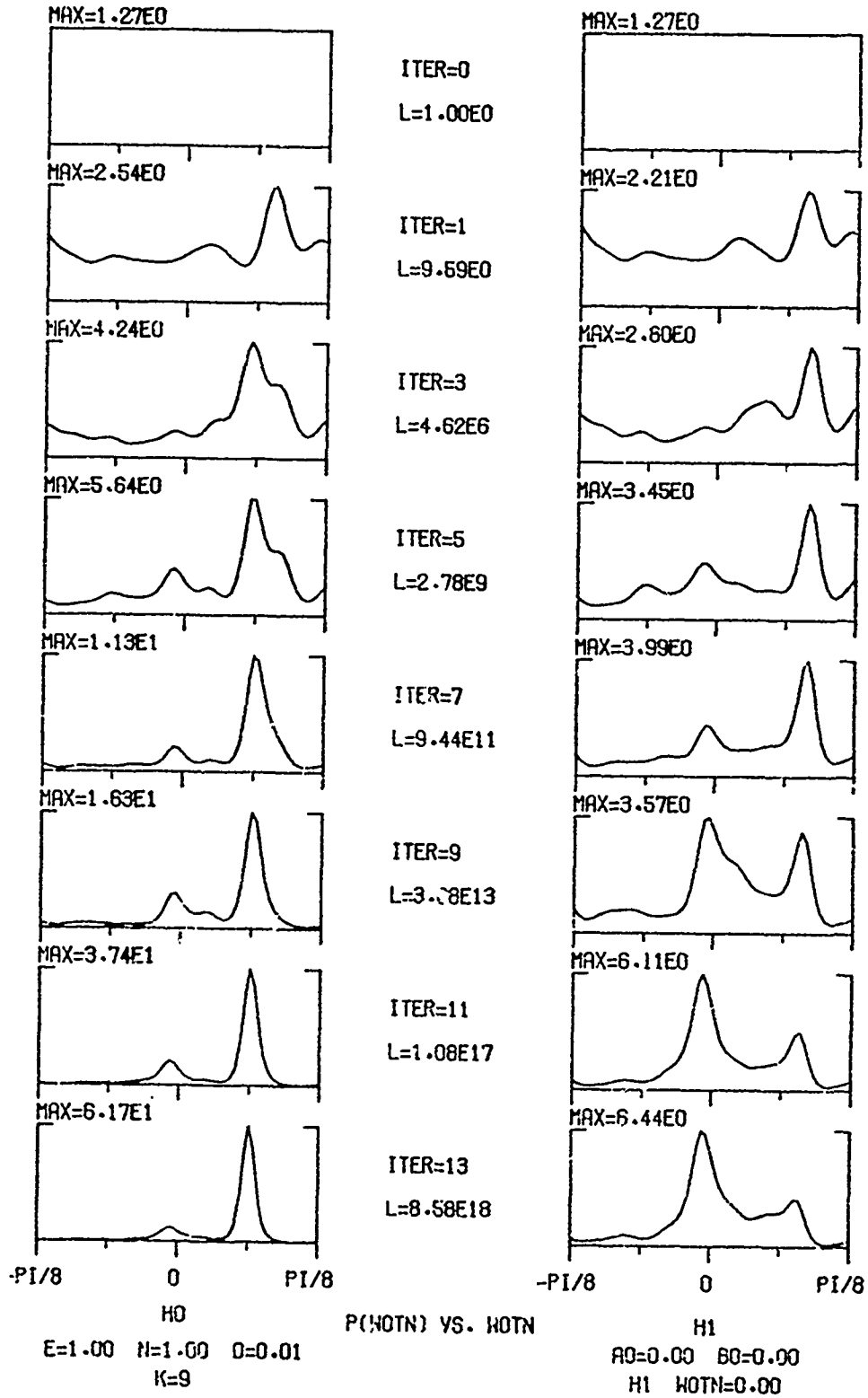


Figure 9.8. Sequential SKE in NUD Simulation.

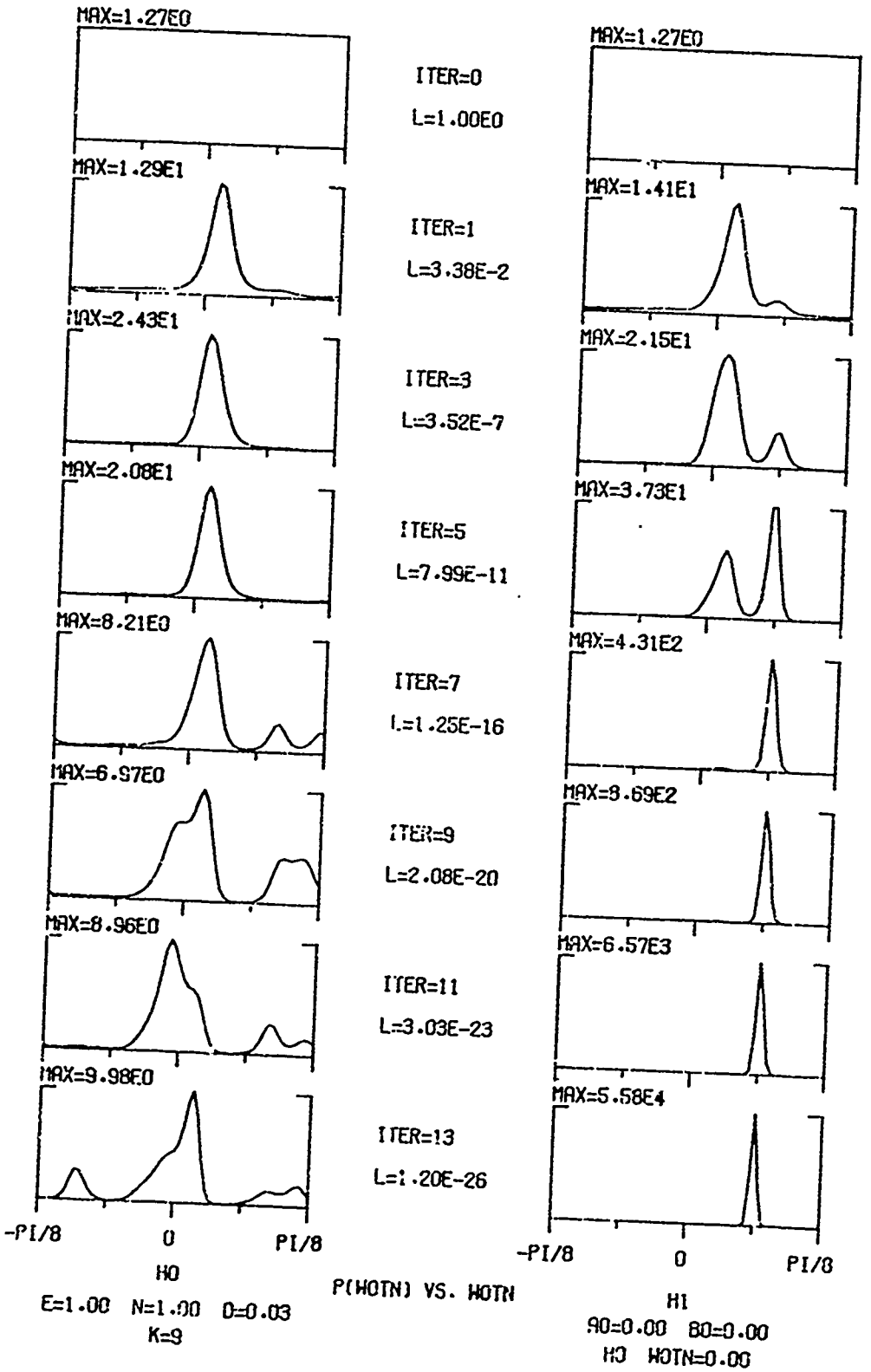


Figure 9.9. Sequential SKE in NUD Simulation.

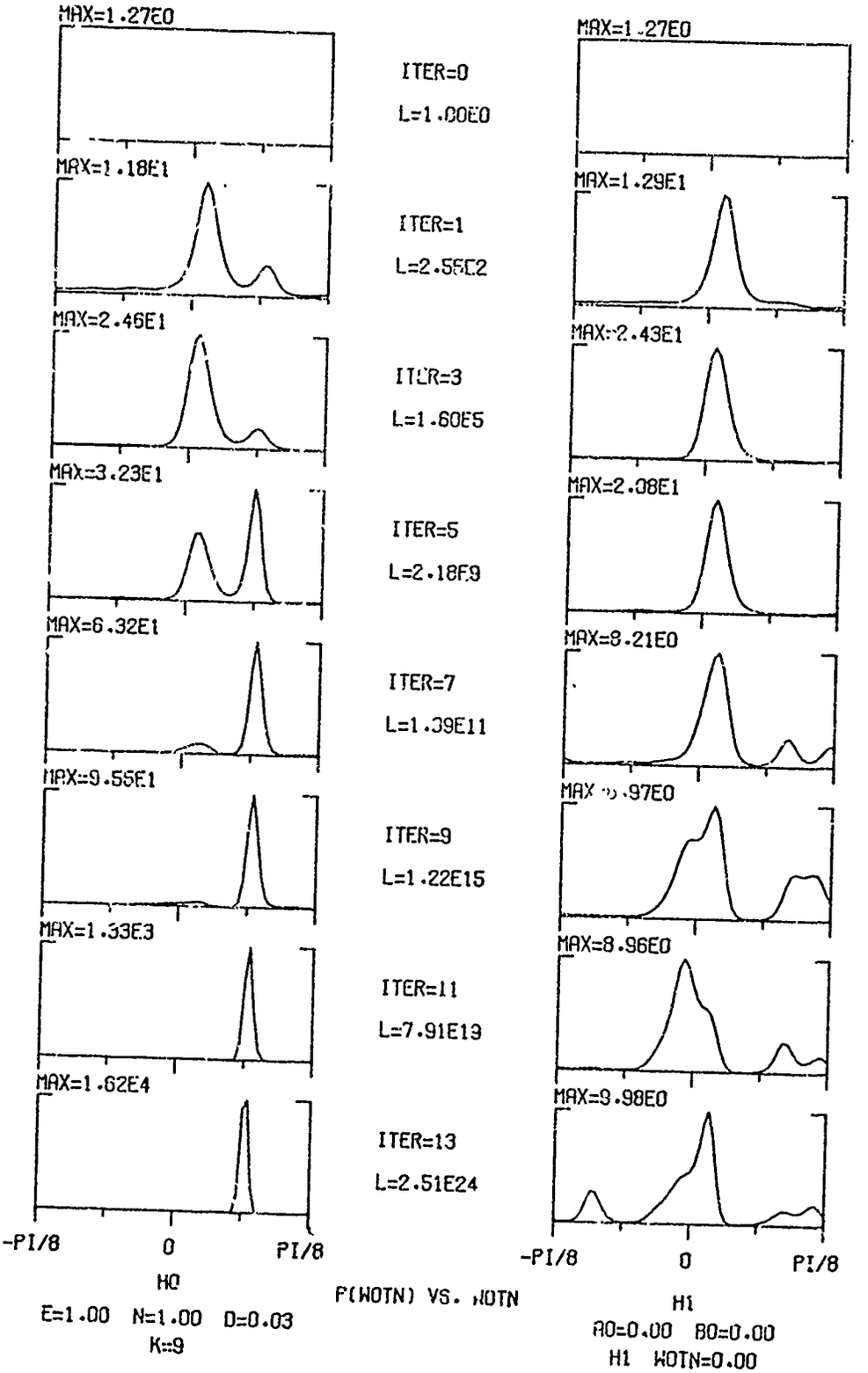


Figure 9.10. Sequential SKE in NUD Simulation.

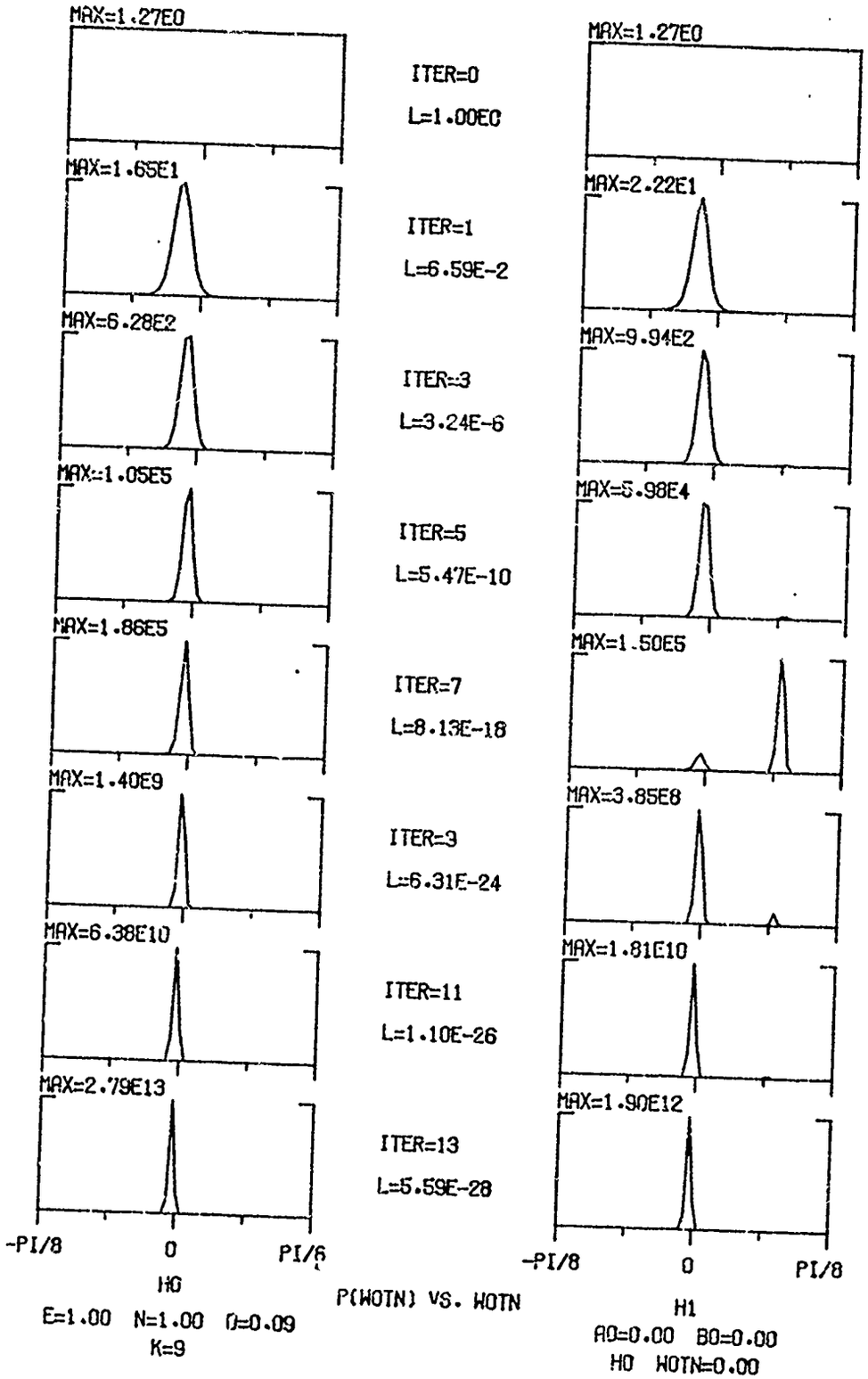


Figure 9.11. Sequential SKE in NUD Simulation.

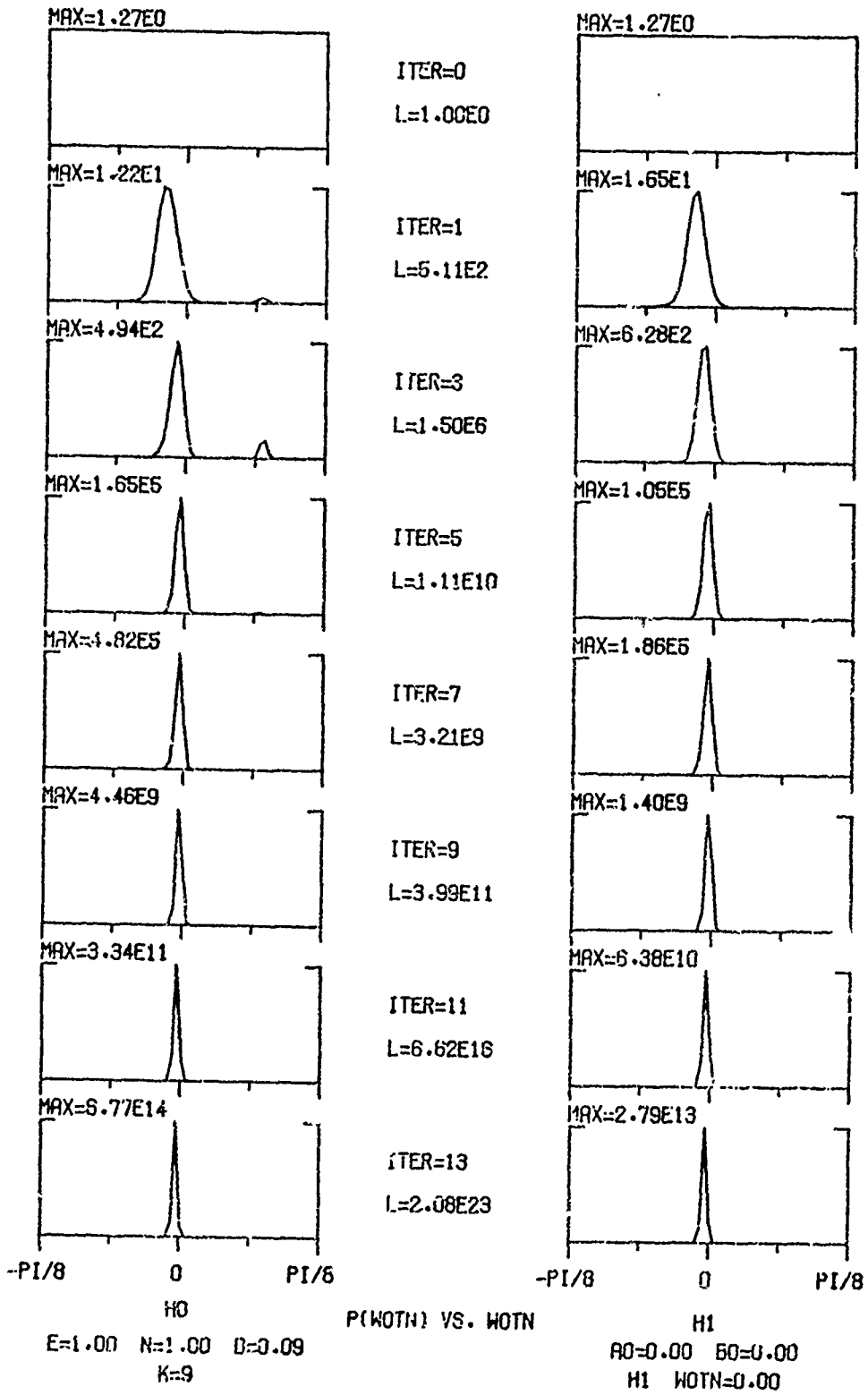


Figure 9.12. Sequential SKE in NUD Simulation.

Chapter X

SUMMARY AND RECOMMENDATIONS FOR FURTHER RESERACH

Summary

This dissertation has taken a global approach to the processing of information from an array of sensors. Essentially, the unprocessed outputs of the individual elements have been considered as the observables. The processor structure was allowed to evolve freely with the sole restriction being the criterion of optimality. Such an approach has been emphasized since it is not clear that the imposition of a structure on an array processor which appears optimal locally (such as the utilization of beamformers, nullformers, and good estimation techniques) will facilitate the overall goal of good signal processing.

Specifically, the array processors discussed were to decide if the random processes observed at the array element outputs consisted of a signal obscured by noise or noise alone. Any uncertain parameters in the problems considered were treated as random variables and knowledge about them was summarized by a priori probability density functions. The resulting detectors were optimum in the sense of making a least-risk decision.

The first task, then, was to specialize the general results of signal detection theory to the optimal processing of data from an array of sensors. The general form of the likelihood ratio was derived based upon observables consisting of the Fourier coefficients of the observed random processes. For a stationary noise field consisting of a componer

independent from sensor to sensor and an additive directional component, the covariance properties of these Fourier coefficients were pursued as a function of the observation period length.

Once the mathematics of the likelihood ratio has been written, the optimal array processor can be implemented in various structures. Four such canonical implementations were discussed: (1) one shot, (2) pseudo estimator, (3) two step, and (4) sequential. The pseudo estimator structure was shown to be the optimal counterpart of a popular ad hoc approach to array processor design where any uncertain parameters are first estimated, then plugged into the parameters known likelihood ratio as if they were known exactly. The general formulation of the time sequential structure revealed that the likelihood ratio can be realized by an appropriate combination of single frequency components. Each is an independent time sequential processor which utilizes its own natural conjugate prior to achieve a certain degree of mathematical tractability.

Of particular interest were three specific problems involving either signal or noise source location uncertainty

(1) Gaussian Signal of Uncertain Direction in Gaussian Noise (GUD)

(2) Signal Known Except for Direction in Noise with an Additive Directional Noise Component of Known Direction (SKED in NKD)

(3) Signal Known Exactly in Noise with an Additive Directional Noise Component of Uncertain Direction (SKE in NUD).

Their likelihood ratios were derived and performance reported for several levels of location uncertainty and two array sizes. Performance was stated in terms of the ROC curve. Several observations were made. It was noted that the optimal array processor for a Gaussian signal of uncertain direction suffered significantly less in performance degradation as

location uncertainty increased than the corresponding processor for a signal of known form. For the signal-to-noise and noise-to-noise ratios investigated, it was found when comparing the SKED in NKD and SKE in NUD processors that uncertainty in the location of a noise source has a less serious effect on performance degradation than a corresponding amount of uncertainty on the location of a known form signal source. Since performance results for two array sizes were investigated, the effect of increasing the array size while keeping the level of uncertainty constant could be studied. When there was signal source location uncertainty, increasing the array size while keeping the level of uncertainty constant led to a degradation in performance. No such performance degradation occurred when the noise source location was uncertain (in fact, a slight increase was noted). Lastly, the performance of a particular class of suboptimal array processors operating in the same environment of uncertainty as their optimal counterparts was investigated. For this class, an estimate and plug structure was imposed with the estimates of signal and noise source location not being data dependent.

Although an estimate and plug structure is appealing due to its explicit adaptive characteristics, it was shown that the optimal array processor exhibits natural learning or adaptive features when implemented sequentially. Computer simulation runs of the SKE in NUD processor were used to illustrate the Bayesian updating which occurs as an integral part of the sequential structure.

Recommendations for Further Research

In the pursuit of a particular piece of research, it is not uncommon for several new and interesting questions to surface. This work is certainly no exception.

Chapter V presented the formulation of the time sequential array processor in general terms. The resulting detector structure illustrated in Figure 5.2 consisted of several independent single frequency components. Within the framework of such a decomposition, it would be interesting to pursue the feasibility of a real-time implementation of the time sequential array processor. Conceivably, each single frequency component would consist of a separate minicomputer dedicated to processing the Fourier coefficients at a single frequency index.

Another question of a more theoretical nature arises in connection with the sequential Bayesian updating of an uncertain parameter's probability density function. Under the correct hypothesis, estimators such as the MLE (maximum likelihood) and MAP (maximum a posteriori) have well known properties. But, what is their nature under the incorrect hypothesis? For example, in Chapter IX it was noted that the a posteriori density of noise source location would often peak up in the direction of the signal source under the hypothesis opposite that actually in force.

Chapters VI and VIII presented detailed derivations of the likelihood ratio and performance results for three specific problems involving either signal or noise source location uncertainty. Their performance was compared to a particular class of suboptimal array processors with an estimate and plug structure to demonstrate the necessity of properly incorporating a priori knowledge into array processor design. The particular class studied was that where the estimates were not data dependent. A valuable extension of those results would be the calculation of performance when data dependent estimators such as the MLE or MAP are used.

APPENDICES

Appendix A

COVARIANCE ARISING FROM A SCALAR RANDOM PROCESS

The following derivation of the covariance between the Fourier coefficients representing a sample function of a scalar random process is taken from Blachman (Blachman, 1957).

Consider the series expansion of a sample function z(t) from a zero mean real random process over the interval (-T/2,T/2)

$$z(t) = \underset{N \to \infty}{\text{l.i.m.}} \sum_{n=-N}^{N} z(n)\left(\frac{1}{T}\right)^{1/2} \exp(jn\omega_0 t) \quad , \quad |t| < T/2 \tag{A.1}$$

where $\omega_0 = \frac{2\pi}{T}$

and $$z(n) = \left(\frac{1}{T}\right)^{1/2} \int_{-T/2}^{T/2} z(t)\exp(-jn\omega_0 t)dt \, . \tag{A.2}$$

Thus, the covariance between any two coefficients can be expressed as

$$E[z(n)z(m)^*] = \frac{1}{T} \int_{-T/2}^{T/2} \int_{-T/2}^{T/2} E[z(u)z(t)]\exp[-j\omega_0(nu-mt)] \, du \, dt \, . \tag{A.3}$$

Making the substitution $t + \tau = u \rightarrow \tau = u - t$

$$E[z(n)z(m)^*] = \frac{1}{T} \int_{-T/2}^{T/2} \int_{-T/2-t}^{T/2-t} E[z(t+\tau)z(t)]\exp\{-j\omega_0[(n-m)t+n\tau]\} \, d\tau \, dt \, . \tag{A.4}$$

Assuming z(t) is stationary (in the wide sense), $E[n(t+\tau)n(t)] = R(\tau)$ and

$$E[z(n)z(m)^*] = \frac{1}{T} \int_{-T/2}^{T/2} \int_{-T/2-t}^{T/2-t} R(\tau)\exp\{-j\omega_0[(n-m)t+n\tau]\} \, d\tau dt \, . \tag{A.5}$$

By the Wiener-Khinchin theorem

$$R(\tau) = \frac{1}{2\pi}\int_{-\infty}^{\infty} N(\omega)\exp(j\omega\tau)d\omega \qquad (A.6)$$

where $N(\omega)$ is the power spectral density function of the random process. Thus, (A.5) becomes

$$E[z(n)z(m)^*] = \frac{1}{2\pi T}\int_{-T/2-T/2-t-\infty}^{T/2\ T/2-t\ \infty} N(\omega)\exp[\omega_0(m-n)t]\exp[j\tau(-n\omega_0+\omega)]\,d\omega\,d\tau\,dt\,. \qquad (A.7)$$

Note that

$$\int_{-T/2-t}^{T/2-t}\exp[j\tau(-n\omega_0+\omega)]d\tau = \frac{1}{j(-n\omega_0+\omega)}\exp[j\tau(-n\omega_0+\omega)]\Big|_{-T/2-t}^{T/2-t}$$

$$= \frac{1}{j(-n\omega_0+\omega)}\exp[-j(-n\omega_0+\omega)t]\{\exp[jT/2(-n\omega_0+\omega)]-\exp[j(-T/2)(-n\omega_0+\omega)]\}$$

$$= \frac{2}{(-n\omega_0+\omega)}\exp[-j(-n\omega_0+\omega)]\sin T/2(-n\omega_0+\omega)\,. \qquad (A.8)$$

After interchanging the order of integration in (A.7)

$$E[z(n)z(m)^*]$$

$$= \frac{1}{\pi T}\int_{-\infty}^{\infty}\int_{-T/2}^{T/2} N(\omega)\,\frac{\sin T/2(-n\omega_0+\omega)}{(-n\omega_0+\omega)}\exp[-j(-n\omega_0+\omega)t]\exp[j\omega_0(m-n)t]\,dt\,d\omega$$

$$= \frac{1}{\pi T}\int_{-\infty}^{\infty}\int_{-T/2}^{T/2} N(\omega)\,\frac{\sin T/2(-n\omega_0+\omega)}{(-n\omega_0+\omega)}\exp[j(m\omega_0-\omega)t]\,dt\,d\omega\,. \qquad (A.9)$$

Using the relationship in (A.8) to integrate over t

$$E[z(n)z(m)^*] = \frac{2}{\pi T}\int_{-\infty}^{\infty} N(\omega)\,\frac{\sin T/2(-n\omega_0+\omega)}{(-n\omega_0+\omega)}\,\frac{\sin T/2(m\omega_0-\omega)}{(m\omega_0-\omega)}\,d\omega\,. \qquad (A.10)$$

Making the substitution $\omega = x\omega_0 \rightarrow x = \omega/\omega_0$

$$E[z(n)z(m)^*] = \frac{2\,\omega_0}{\pi T} \int_{-\infty}^{\infty} N(x\omega_0) \frac{\sin \pi/\omega_0 (-n\omega_0+x\omega_0)}{(-n\omega_0+x\omega_0)} \frac{\sin \pi/\omega_0 (m\omega_0-x\omega_0)}{(m\omega_0-x\omega_0)}\, dx$$

$$= \int_{-\infty}^{\infty} N(x\omega_0) \frac{\sin \pi\,(x-n)}{\pi(x-n)} \frac{\sin \pi\,(x-m)}{\pi(x-m)}\, dx \quad . \qquad (A.11)$$

The covariance between any two Fourier coefficients now can be written as

$$E[z(n)z(m)^*] = \int_{-\infty}^{\infty} N(x\omega_0)\, \mathrm{sinc}(x-n)\, \mathrm{sinc}(x-m)\, dx \qquad (A.12)$$

where $$\mathrm{sinc}(x) = \frac{\sin \pi x}{\pi x} \quad .$$

Note the orthonormal property of sinc(x)

$$\int_{-\infty}^{\infty} \mathrm{sinc}(x-n)\, \mathrm{sinc}(x-m)\, dx = \delta_{nm} \qquad (A.13)$$

where $$\delta_{nm} = \begin{cases} 1\,, & n = m \\ 0\,, & n \neq m \end{cases} \quad .$$

Appendix B

COVARIANCE MATRICIES FOR BANDLIMITED WHITE GAUSSIAN NOISE

This section contains a set of five covariance matricies (Figures B.1-B.5) which represent successive doubling of the observation period length (Marshall, 1973). The Fourier coefficients representing a sample function from the scalar random process are related by the expression

$$E[z(n)z(m)^*] = \int_{-\infty}^{\infty} N(x\omega_0)\text{sinc}(x-n)\text{sinc}(x-m)dx \qquad (B.1)$$

where $\omega_0 = \frac{2\pi}{T}$

$N(x\omega_0)$ = power spectral density function of the random process

and $\text{sinc}(x) = \frac{\sin(\pi x)}{(\pi x)}$.

The spectrum considered is that of unit height white Gaussian noise bandlimited to W Hertz. In the first covariance matrix, the observation length is such that $W = .5\omega_0/2\pi$ or $T = .5W$. In the fifth covariance matrix, the observation length is sixteen times longer (i.e., $W = 8\omega_0/2\pi$ or $T = 8/W$). Note how the covariance matricies become progressively more diagonal in form as T is increased.

n \ m	-6	-5	-4	-3	-2	-1	0	1	2	3	4	5	6
-6	.00	.00	.00	.00	.00	-.01	.00	.01	.00	.00	.00	.00	.00
-5	.00	.00	.00	.00	-.01	.01	.00	-.01	.01	.00	.00	.00	.00
-4	.00	.00	.00	.00	.01	-.02	.00	.01	-.01	.00	.00	.00	.00
-3	.00	.00	.00	.01	-.01	.02	.01	-.02	.01	-.01	.00	.00	.00
-2	.00	-.01	.01	-.01	.01	-.02	-.01	.03	-.01	.01	-.01	.01	.00
-1	-.01	.01	-.02	.02	-.03	.03	.06	-.06	.03	-.02	.01	-.01	.01
0	.00	.00	.00	.01	-.01	.06	.77	.06	-.01	.01	.00	.00	.00
1	.01	-.01	.01	-.02	.03	-.06	.06	.08	-.03	.02	-.02	.01	-.01
2	.00	.01	-.01	.01	-.01	.03	-.01	-.03	.01	-.01	.01	-.01	.00
3	.00	.00	.00	-.01	.01	-.02	.01	.02	-.01	.01	.00	.00	.00
4	.00	.00	.00	.00	-.01	.01	.00	-.02	.01	.00	.00	.00	.00
5	.00	.00	.00	.00	.01	-.01	.00	.01	-.01	.00	.00	.00	.00
6	.00	.00	.00	.00	.00	.01	.00	-.01	.00	.00	.00	.00	.00

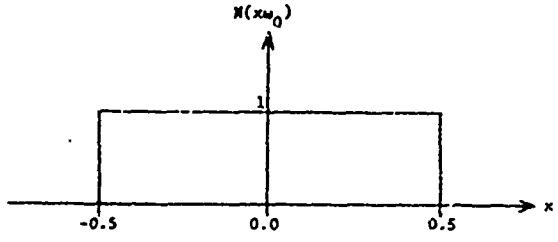


Figure B.1. Covariance Matrix No. 1.

n \ m	-6	-5	-4	-3	-2	-1	0	1	2	3	4	5	6
-6	.00	.00	.00	.00	.01	-.03	.00	.02	-.01	.00	.00	.00	.00
-5	.00	.00	.00	.00	-.01	.03	.00	-.03	.01	.00	.00	.00	.00
-4	.00	.00	.00	-.01	.01	-.04	-.01	.04	-.01	.00	.00	.00	.00
-3	.00	.00	-.01	.01	-.02	.06	.01	-.05	.02	.00	.00	.00	.00
-2	.01	-.01	.01	-.02	.03	-.10	-.03	.07	-.03	.02	-.01	.01	-.01
-1	-.03	.03	-.04	.06	-.10	.47	.16	-.16	.07	-.05	.04	-.03	.02
0	.00	.00	-.01	.01	-.03	.16	.90	.16	-.03	.01	-.01	.00	.00
1	.02	-.03	.04	-.05	.07	-.16	.16	.47	-.10	.06	-.04	.03	-.02
2	-.01	.01	-.01	.02	-.03	.07	-.03	-.10	.03	-.02	.01	-.01	.01
3	.00	.00	.00	.00	.02	-.05	.01	.06	-.02	.01	-.01	.00	.00
4	.00	.00	.00	.00	-.01	.04	-.01	-.04	.01	-.01	.00	.00	.00
5	.00	.00	.00	.00	.01	-.03	.00	.03	-.01	.00	.00	.00	.00
6	.00	.00	.00	.00	-.01	.02	.00	-.03	.01	.00	.00	.00	.00

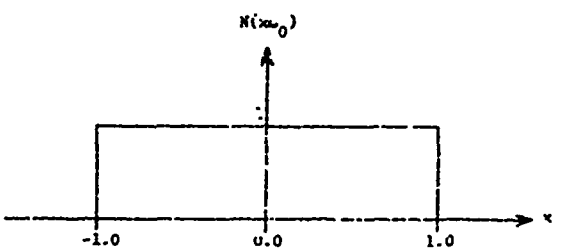


Figure B.2. Covariance Matrix No. 2.

n	-6	-5	-4	-3	-2	-1	0	1	2	3	4	5	6
-6	.00	.00	.00	.00	.04	.00	-.01	.01	-.03	.00	.00	.00	.00
-5	.00	.00	.00	.02	-.05	.00	.01	-.02	.03	.00	.00	.00	.00
-4	.00	.00	.00	-.03	.07	.00	-.01	.02	-.04	.00	.00	.00	.00
-3	.00	.02	-.03	.04	-.11	-.01	.03	-.03	.05	.00	.00	.00	.00
-2	.04	-.05	.07	-.11	.49	.14	-.10	.08	-.10	.05	-.04	.03	-.03
-1	.00	.00	.00	-.01	.14	.93	.05	.05	.08	-.03	.02	-.02	.01
0	-.01	.01	-.01	.03	-.10	.05	.95	.05	-.10	.03	-.01	.01	-.01
1	.01	-.02	.02	-.03	.08	.05	.05	.93	.14	-.01	.00	.00	.00
2	.03	.03	-.04	.05	-.10	.08	-.10	.14	.49	-.11	.07	-.05	.04
3	.00	00	.00	.00	.05	-.03	.03	-.01	-.11	.04	-.03	.02	.00
4	.00	.00	.00	.00	-.04	.02	-.01	.00	.07	-.03	.00	.00	.00
5	.00	.00	.00	.00	.03	-.02	.01	.00	-.05	.02	.00	.00	.00
6	.00	.00	.00	.00	-.03	.01	-.01	.00	.04	.00	.00	.00	.00

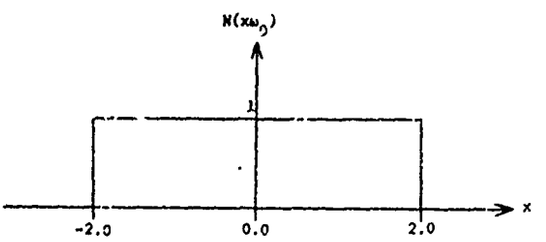


Figure B.3. Covariance Matrix No. 3.

n \ m	-9	-8	-7	-6	-5	-4	-3	-2	-1	0	1	2	3	4	5	6	7	8	9
-9	.00	.00	.00	.00	.00	-.04	.01	.00	.00	.01	-.01	.01	-.01	.02	.00	.00	.00	.00	.00
-8	.00	.00	.00	.00	.00	.04	-.01	.00	.00	-.01	.01	-.01	.01	-.02	.00	.00	.00	.00	.00
-7	.00	.00	.00	.00	.02	-.05	.01	.00	-.01	.01	-.01	.01	-.02	.03	.00	.00	.00	.00	.00
-6	.00	.00	.00	.00	-.03	.07	-.01	-.01	.01	-.01	.02	-.02	.02	-.03	.00	.00	.00	.00	.00
-5	.00	.00	.02	-.03	.06	-.12	-.01	.02	-.02	.02	-.02	.02	-.03	.04	.00	.00	.00	.00	.00
-4	-.04	.04	-.05	.07	-.12	.49	.13	-.04	.07	-.06	.05	-.05	.05	-.06	.04	-.03	.03	-.02	.02
-3	.01	-.01	.01	-.01	-.01	.13	.94	.00	-.04	.03	-.03	.03	-.03	.05	-.03	.02	-.02	.01	-.01
-2	.00	.00	.00	-.01	.02	-.09	.04	.97	.03	-.03	.03	-.03	.03	-.05	.02	-.02	.01	-.01	.01
-1	.00	.00	-.01	.01	-.02	.07	-.04	.03	.97	.03	-.03	.03	-.03	.05	-.02	.02	-.01	.01	-.01
0	.01	-.01	.01	-.01	.02	-.06	.03	-.03	.03	.97	.03	-.03	.03	-.06	.02	-.01	.01	-.01	.01
1	-.01	.01	-.01	.02	-.02	.05	.03	.03	-.03	.03	.97	.03	-.04	.07	-.02	.01	-.01	.00	.00
2	.01	-.01	.01	-.02	.02	-.05	.03	-.03	.03	-.03	.03	.97	.04	-.09	.02	-.01	.00	.00	.00
3	-.01	.01	-.02	.02	-.03	.05	-.03	.03	-.03	.03	-.04	.04	.94	.13	-.01	-.01	.01	-.01	.01
4	0.2	-.02	.03	-.03	.04	-.06	.05	-.05	.05	-.06	.07	-.09	.13	.49	-.12	.07	-.05	.04	-.04
5	.00	.00	.00	.00	.00	.04	-.03	.02	-.02	.02	-.02	.02	-.01	-.12	.04	-.03	.02	.00	.00
6	.00	.00	.00	.00	.00	-.03	.02	-.02	.02	-.01	.01	-.01	-.01	.07	-.03	.00	.00	.00	.00
7	.00	.00	.00	.00	.00	.03	-.02	.01	-.01	.01	-.01	.00	.01	-.05	.02	.00	.00	.00	.00
8	.00	.00	.00	.00	.00	-.02	.01	-.01	.01	-.01	.00	.00	-.01	.04	.00	.00	.00	.00	.00
9	.00	.00	.00	.00	.00	.02	-.01	.01	-.01	.01	.00	.00	.01	-.04	.00	.00	.00	.00	.00

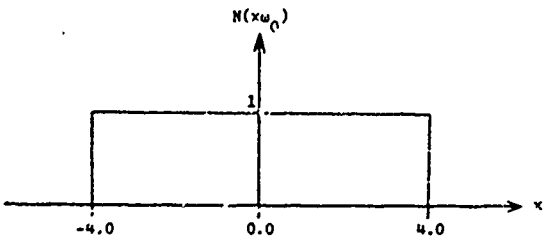


Figure .4. Covariance Matrix No. 4.

m \ n	-9	-8	-7	-6	-5	-4	-3	-2	-1	0	1	2	3	4	5	6	7	8	9
-9	.05	-.12	.03	.01	-.02	.02	-.02	.02	-.02	.02	-.02	.02	.00	.02	-.02	.00	-.02	.02	.06
-8	-.12	.49	.13	-.04	.05	-.05	.04	-.04	.04	-.04	.05	-.05	.04	-.04	.03	-.03	.03	-.04	.02
-7	.03	.01	.95	.04	-.04	.03	-.03	.02	-.02	.02	.00	.02	-.02	.04	-.02	.02	-.02	.03	-.02
-6	.01	-.04	.04	.97	.02	-.03	.02	-.02	.02	-.02	.02	-.02	.02	-.02	.02	-.02	.02	-.03	.00
-5	-.02	.06	-.04	.02	.98	.02	-.02	.02	-.02	.01	.01	.02	-.01	.01	-.01	.02	-.02	.03	-.02
-4	.02	-.05	.03	-.02	.02	.98	.02	-.01	.02	-.01	.01	-.01	.01	-.01	.01	-.02	.02	-.03	.02
-3	-.02	.04	-.03	.02	-.02	.02	.98	.01	-.01	.01	-.01	.01	-.01	.01	-.01	.02	-.02	.03	.00
-2	.02	-.04	.02	-.02	.02	-.01	.01	.99	.01	-.01	.01	-.01	.01	-.01	.01	-.02	.02	-.03	.02
-1	-.02	.04	-.02	.02	-.02	.01	-.01	.01	.99	.01	-.01	.01	-.01	.01	-.01	.02	.00	.03	-.02
0	.02	-.03	.02	-.02	.01	-.01	.01	-.01	.01	.99	.01	-.01	.01	-.01	.01	-.02	.02	-.03	.02
1	-.02	.03	.00	.02	-.01	.01	-.01	.01	-.01	.01	.99	.01	-.01	.01	-.02	.02	-.02	.04	-.02
2	.02	-.03	.02	-.02	.01	-.01	.01	-.01	.00	-.01	.01	.99	.01	-.01	.02	-.02	.02	-.04	.02
3	.00	.03	-.02	.02	-.01	.01	-.01	.01	.01	.01	-.01	.01	.98	.02	-.02	.02	-.03	.04	-.02
4	.02	-.03	.02	-.02	.01	-.01	.02	-.01	.01	-.01	.02	-.01	.02	.98	.02	-.02	.03	-.05	.02
5	-.02	.03	-.02	.02	-.01	.01	-.01	.01	-.01	.01	-.02	.02	-.02	.02	.98	.02	-.03	.06	-.02
6	.00	-.03	.02	-.02	.02	-.02	.02	-.02	.02	-.02	.02	-.02	.02	-.02	.02	.97	.04	-.08	.01
7	-.02	.03	-.02	.02	-.02	.02	-.02	.02	.00	.02	-.02	.02	-.03	.03	-.03	.04	.95	.13	.00
8	.02	-.03	.03	-.03	.03	-.03	.03	-.03	.03	-.03	.04	-.04	.04	-.05	.06	-.08	.01	.49	-.12
9	.02	.02	.02	.00	-.02	.02	.00	.02	-.02	.02	-.02	.02	-.02	.02	-.02	.01	.00	-.12	.05

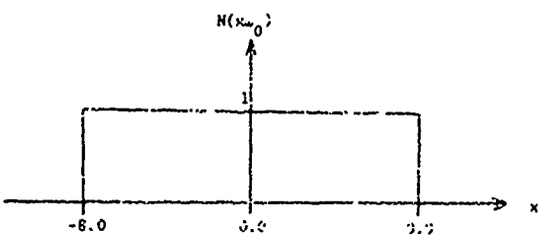


Figure B.5. Covariance Matrix No. 5.

Appendix C

COVARIANCE ARISING FROM A VECTOR RANDOM PROCESS

Consider the vector of sample functions $\underline{z}(t)$ from the zero mean stationary vector random process observed as the collection of outputs from all the array elements. Appendix A has already derived the relationship between the Fourier coefficients representing a single element's output. When the noise field contains an additive directional noise component, the Fourier coefficients arising from the outputs of two different array elements will become related since a portion of the random process observed at one element simply will be a time delayed version of that observed at another. Let

$$z_\ell(t) = d_\ell(t)$$

and

$$z_k(t) = d_\ell(t-\tau_{\ell k}) \tag{C.1}$$

where $\tau_{\ell k}$ is the time delay between the ℓ^{th} and k^{th} elements of the directional noise component $d_\ell(t)$. The Fourier coefficients representing each observed waveform are

$$z_\ell(n) = \left(\frac{1}{T}\right)^{1/2} \int_{-T/2}^{T/2} d_\ell(t)\exp(-jn\omega_0 t)\, dt \tag{C.2}$$

$$z_k(m) = \left(\frac{1}{T}\right)^{1/2} \int_{-T/2}^{T/2} d_\ell(t-\tau_{\ell k})\exp(-jm\omega_0 t)\, dt$$

$$= \exp(-jm\omega_0\tau_{\ell k}) \cdot \left(\frac{1}{T}\right)^{1/2} \int_{-T/2-\tau_{\ell k}}^{T/2-\tau_{\ell k}} d_\ell(t)\exp(-jm\omega_0 t)\, dt\,. \tag{C.3}$$

Thus, the covariance between these two coefficients can be expressed as

$$E[z_\ell(n)z_k(m)^*|\tau_{\ell k}] = \exp(-jm\omega_0\tau_{\ell k})$$

$$\cdot \frac{1}{T}\int_{-T/2-\tau_{\ell k}}^{T/2-\tau_{\ell k}} \int_{-T/2}^{T/2} E[d_\ell(u)d_\ell(t)]\exp[-j\omega_0(nu-mt)]\,du\,dt\,. \tag{C.4}$$

The remainder of the derivation is similar to that pursued in Appendix A with the following exceptions

(1) (A.6) becomes

$$R(\tau) = \frac{1}{2\pi}\int_{-\infty}^{\infty} D(\omega)\exp(j\omega\tau)\,d\omega \tag{C.5}$$

where $D(\omega)$ is the power spectral density function of the directional noise

(2) (A.9) and (A.10) become

$$E[z_\ell(n)z_k(m)^*|\tau_{\ell k}] = \exp(-jm\omega_0\tau_{\ell k})$$

$$\cdot \frac{1}{\pi T}\int_{\infty}^{\infty} \int_{-T/2-\tau_{\ell k}}^{T/2-\tau_{\ell k}} D(\omega)\,\frac{\sin T/2(-n\omega_0+\omega)}{(-n\omega_0+\omega)}\exp[j(m\omega_0-\omega)t]\,dt\,d\omega$$

$$= \exp(-jm\omega_0\tau_{\ell k})$$

$$\cdot \frac{2}{\pi T}\int_{-\infty}^{\infty} D(\omega)\,\frac{\sin T/2(-n\omega_0+\omega)}{(-n\omega_0+\omega)}\,\frac{\sin T/2(m\omega_0-\omega)}{(m\omega_0-\omega)}\exp[-j(m\omega_0-\omega)\tau_{\ell k}]\,d\omega \tag{C.6}$$

and (3) (A.12) becomes

$$E[z_\ell(n)z_k(m)^*|\tau_{\ell k}] = \exp(-jm\omega_0\tau_{\ell k})$$

$$\cdot \int_{-\infty}^{\infty} D(x\omega_0)\mathrm{sinc}(x-n)\mathrm{sinc}(x-m)\exp[-j(x-m)(-\frac{\tau_{\ell k}}{T})2\pi]\,dx. \tag{C.7}$$

An interesting special case of (C.7) occurs when $D(x\omega_0)$ is white (i.e., $D(x\omega_0)$ is not a function of $x\omega_0$) and n = m. Setting $D(x\omega_0) = D$

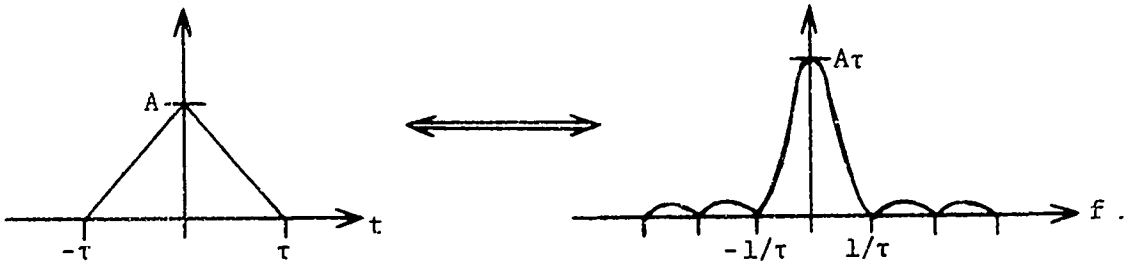
$$E[z_\ell(n)z_k(n)^*|\tau_{\ell k}] = \exp(-jn\omega_0\tau_{\ell k}) \cdot D \int_{-\infty}^{\infty} \text{sinc}^2(x-n)\exp[-j(x-n)(-\frac{\tau_{\ell k}}{T})2\pi]dx$$

$$= \exp(-jn\omega_0\tau_{\ell k}) \cdot D \int_{-\infty}^{\infty} \text{sinc}^2(f)\exp[j2\pi ft]df \qquad (C.8)$$

where $f = x - n$ and $t = \tau_{\ell k}/T$.

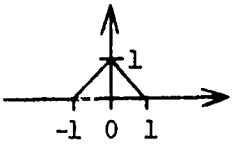
Note the Fourier transform pair

$$A\Lambda(t/\tau) \Longleftrightarrow A\tau \ \text{sinc}^2 f\tau \qquad (C.9)$$



Thus

$$E[z_\ell(n)z_k(n)^*|\tau_{\ell k}] = \exp(-jn\omega_0\tau_{\ell k}) \cdot D \cdot \{$$



$$\}. \qquad (C.10)$$

Appendix D

COVARIANCE BETWEEN ADJACENT OBSERVATION VECTORS

The total observation period now will be broken into several smaller incremental periods. Thus, the vector of sample functions observed at the outputs of all the array elements will be represented by a sequence of observation vectors $\underline{z}^i$. In this section, the covariance between the Fourier coefficients of adjacent observation vectors will be derived. Assume the noise field consists of a portion which is independent from sensor to sensor plus an additive directional noise component. Let

$$z_\ell^i(t) = n_\ell^i(t) + d_\ell^i(t)$$

and

$$\begin{aligned} z_k^{i+1} &= n_k^{i+1}(t) + d_\ell^{i+1}(t-\tau_{\ell k}) \\ &= n_k^i(t+T) + d_\ell^i(t+T-\tau_{\ell k}) \end{aligned} \tag{D.1}$$

where $n^i(t)$ and $n_k^{i+1}(t)$ are the portions of the observed waveforms due to the independent noise, $d_\ell^i(t)$ and $d_\ell^{i+1}(t-\tau_{\ell k})$ are due to the directional noise component, and $\tau_{\ell k}$ is the time delay between the ℓ^{th} and k^{th} elements of the directional noise. The Fourier coefficients representing each observed waveform are

$$z_\ell^i(n) = \left(\frac{1}{T}\right)^{1/2} \int_{-T/2}^{T/2} [n_\ell^i(t) + d_\ell^i(t)] \exp(-jn\omega_0 t)\, dt \tag{D.2}$$

$$z_k^{i+1}(m) = \left(\frac{1}{T}\right)^{1/2} \int_{-T/2}^{T/2} [n_k^i(t+T) + d_\ell^i(t+T-\tau_{\ell k})] \exp(-jm\omega_0 t)\, dt$$

$$= \exp[-jm\omega_0 T] \cdot \left(\frac{1}{T}\right)^{1/2} \int_{T/2}^{3T/2} n_k^i(t) \exp(-jm\omega_0 t)\, dt$$

$$+ \exp[-jm\omega_0 (T-\tau_{\ell k})] \cdot \left(\frac{1}{T}\right)^{1/2} \int_{T/2-\tau_{\ell k}}^{3T/2-\tau_{\ell k}} d_\ell^i(t) \exp(-jm\omega_0 t)\, dt \,. \tag{D.3}$$

Assuming the independent and directional noise are independent, the covariance between these two coefficients can be expressed as

$$E[z_\ell^i(n) z_k^{i+1}(m)^* | \tau_{\ell k}] = \exp[-jm\omega_0 (T-\tau_{\ell k})]$$

$$\cdot \left(\frac{1}{T}\right)^{1/2} \int_{T/2-\tau_{\ell k}}^{3T/2-\tau_{\ell k}} \int_{-T/2}^{T/2} E[n_\ell^i(t) n_k^i(u) + d_\ell^i(t) d_k^i(u)]$$

$$\cdot \exp[-j\omega_0 (nu-mt)]\, du\, dt \,. \tag{D.4}$$

Note that $E[n_\ell^i(t) n_k^i(u)] = 0 \,, \quad \ell \neq k \,.$

The remainder of the derivation is similar to that pursued in Appendix C with the following exceptions

(1) (C.5) becomes

$$R(\tau) = \frac{1}{2\pi} \int_{-\infty}^{\infty} [N(\omega) + D(\omega)] \exp(j\omega t)\, d\omega \tag{D.5}$$

(2) (C.6) becomes

$$E[z_\ell^i(n) z_k^{i+1}(m)^* | \tau_{\ell k}] = \exp[-jm\omega_0 (T-\tau_{\ell k})]$$

$$\cdot \frac{2}{\pi T} \int_{-\infty}^{\infty} [N(\omega) \cdot \delta_{\ell k} + D(\omega)] \frac{\sin T/2(-n\omega_0+\omega)}{(-n\omega_0+\omega)} \frac{\sin T/2(m\omega_0-\omega)}{(m\omega_0-\omega)} \exp[-j(\omega - m\omega_0 (t-\tau_{\ell k})]\, d\omega \tag{D.6}$$

where $\delta_{\ell k} = \begin{cases} 1 \,, & \ell = k \\ 0 \,, & \ell \neq k \end{cases}$

and (3) (C.7) becomes

$$E[z_\ell^i(n)z_k^{i+1}(m)^*|\tau_{\ell k}] = \exp[-jm\omega_0(T-\tau_{\ell k})]$$

$$\cdot \int_{-\infty}^{\infty} [N(x\omega_0)\cdot\delta_{\ell k}+D(x\omega_0)]\text{sinc}(x-n)\text{sinc}(x-m)\exp[-j(x-m)(1-\frac{\tau_{\ell k}}{T})2\pi]\,dx\ . \tag{D.7}$$

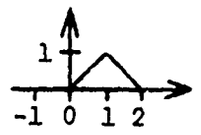
In a fashion similar to Appendix C, an interesting special case of (D.7) occurs when both $N(x\omega_0)$ and $D(x\omega_0)$ are white and $n = m$. Setting $N(x\omega_0) = N$ and $D(x\omega_0) = D$

$$\begin{aligned} E[z_\ell^i(n)z_k^{i+1}(n)^*|\tau_{\ell k}] &= \exp[-jn\omega_0(T-\tau_{\ell k})]\cdot [N\cdot\delta_{\ell k}+D] \\ &\quad \cdot \int_{-\infty}^{\infty} \text{sinc}^2(x-n)\exp[-j(x-n)(1-\frac{\tau_{\ell k}}{T})2\pi]\,dx \\ &= \exp[-jn\omega_0(T-\tau_{\ell k}) \cdot [N\cdot\delta_{\ell k}+D] \\ &\quad \cdot \int_{-\infty}^{\infty} \text{sinc}^2(f)\exp(j2\pi ft)\,df \end{aligned} \tag{D.8}$$

where $f = x - n$ and $t = \tau_{\ell k}/T - 1$.

Thus, making use of the Fourier transform pair in (C.9)

$$E[z_\ell^i(n)z_k^{i+1}(n)^*|\tau_{\ell k}] = \exp[-jn\omega_0(T-\tau_{\ell k})] \cdot [N\cdot\delta_{\ell k}+D] \cdot \{$$



$\}$. (D.9)

Appendix E

MATRIX INVERSION

The covariance matricies encountered in the problems discussed in Chapter VI can be inverted using a technique due to Bartlett (Bartlett, 1951). Let the matrix to be inverted be of the form

$$\underline{A}_0 = \underline{B} + \underline{c}\,\underline{d}^* \tag{E.1}$$

where $\underline{A}_0$ and $\underline{B}$ are K x K matrices and $\underline{c}$ and $\underline{d}$ are K-dimensional column vectors. Proceeding formally

$$\begin{aligned}\underline{A}_0^{-1} &= (\underline{B} + \underline{c}\,\underline{d}^*)^{-1} \\ &= \underline{B}^{-1}\{1 - \underline{c}\,\underline{d}^*\underline{B}^{-1} + (\underline{c}\,\underline{d}^*\underline{B}^{-1})^2 - \ldots\} \\ &= \underline{B}^{-1} - \underline{B}^{-1}\,\underline{c}\,\underline{d}^*\underline{B}^{-1}\{1 - \underline{d}^*\underline{B}^{-1}\underline{c} + (\underline{d}^*\underline{B}^{-1}\underline{c})^2 - \ldots\}\end{aligned}$$

where we note that $\underline{d}^*\underline{B}^{-1}\underline{c}$ is a complex scalar. Summing the geometric series completes the derivation

$$\underline{A}_0^{-1} = \underline{B}^{-1} - \frac{\underline{B}^{-1}\,\underline{c}\,\underline{d}^*\underline{B}^{-1}}{1 + \underline{d}^*\underline{B}^{-1}\underline{c}} . \tag{E.2}$$

This result can be extended as follows. Let the matrix to be inverted be of the form

$$\begin{aligned}\underline{A}_1 &= \underline{B} + \underline{c}\,\underline{d}^* + \underline{e}\,\underline{f}^* \\ &= \underline{A}_0 + \underline{e}\,\underline{f}^*\end{aligned} \tag{E.3}$$

where $\underline{e}$ and $\underline{f}$ are K-dimensional column vectors. Thus, utilizing (E.2)

$$\underline{A}_1^{-1} = \underline{A}_0^{-1} - \frac{\underline{A}_0^{-1}\,\underline{e}\,\underline{f}^*\,\underline{A}_0^{-1}}{1 + \underline{f}^*\underline{A}_0^{-1}\,\underline{e}} . \tag{E.4}$$

Appendix F

THE CALCULATION OF PERFORMANCE FOR LIKELIHOOD RATIO PROCESSORS VIA COMPUTER SIMULATION*

A complete description of an optimum detector includes not only the likelihood ratio Λ but also the processor's ROC(receiver operating characteristic) curves. The false alarm probability P_F and detection probability P_D are

$$P_F = \int_{\eta}^{\infty} p(\Lambda|H_0)\, d\Lambda \tag{F.1}$$

and

$$P_D = \int_{\eta}^{\infty} p(\Lambda|H_1)\, d\Lambda \tag{F.2}$$

where η is the detection threshold value and $p(\Lambda|H_0)$ and $p(\Lambda|H_1)$ are the probability density functions of Λ under the two mutually exclusive and exhaustive hypotheses H_1 that a signal is present and H_0 that it is not. Peterson, Birdsall, and Fox (Peterson, Birdsall, and Fox, 1954) have shown that

$$p(\Lambda|H_1) = \Lambda p(\Lambda|H_0)\,. \tag{F.3}$$

Thus, P_D may be written equivalently as

$$P_D = \int_{\eta}^{\infty} \Lambda p(\Lambda|H_0)\, d\Lambda\,. \tag{F.4}$$

The expression (F.4) is particularly valuable when the densities of Λ under the two hypotheses cannot be determined analytically. Typically, one

---

* The text of this appendix has been accepted for publication in the IEEE Trans. on Information Theory (Hodgkiss and Nolte, 1975).

then carries out a Monte Carlo simulation of the optimum processor and from his results forms estimates $\hat{p}(\Lambda|H_0)$ and $\hat{p}(\Lambda|H_1)$ of the desired densities. We see that (F.4) implies that only the density under H_0 actually need be obtained. One clear benefit of this approach is eliminating the need to simulate the signal (which may contain uncertain parameters) along with the noise to provide receiver input under H_1. Here we make the distinction between the optimum processor which contains both signal and noise parameters and the simulated input to the processor which may consist of noise alone or signal plus noise depending on whether the true hypothesis is H_0 or H_1.

Note that (F.4) places a great deal of emphasis on the upper tail of $p(\Lambda|H_0)$. An equivalent expression is

$$P_D = 1 - \int_0^\eta \Lambda p(\Lambda|H_0)\, d\Lambda . \qquad (F.5)$$

For many problems, particularly those whose underlying statistics are Gaussian, $p(\Lambda|H_0)$ will have both a large portion of its probability mass concentrated in the region $\Lambda \leq 1$ and a long, slowly decaying tail extending to infinity. We would expect any computer simulation to generate only a few observations in this extended tail region. Any calculation which emphasizes these points (such as the first moment) will be strongly affected by their exact location. In this case, the emphasis arises out of the multiplication of $p(\Lambda|H_0)$ by Λ. Thus, (F.5) is proposed as a more desirable means of calculating P_D than (F.4) from the standpoint of utilizing numerical simulation results since it avoids heavily weighting the tail region as much as possible. A more quantitative rationale is discussed in the appendix. Summarizing the relevant equations for estimating the ROC we have

$$\hat{P}_F = \int_\eta^\infty \hat{p}(\Lambda|H_0)\, d\Lambda \tag{F.6}$$

$$\hat{P}_D = 1 - \int_0^\eta \Lambda\hat{p}(\Lambda|H_0)\, d\Lambda \tag{F.7}$$

where the circumflexes denote estimates of the true values.

If we had been unaware of the long upper tail, instead of (F.7) in calculating the ROC we might have used

$$\hat{P}_D' = \int_\eta^\infty \Lambda\hat{p}(\Lambda|H_0)\, d\Lambda\,. \tag{F.8}$$

Note that (F.7) may be rewritten as

$$\begin{aligned}\hat{P}_D &= 1 - \int_0^\infty \Lambda\hat{p}(\Lambda|H_0)\, d\Lambda + \int_\eta^\infty \Lambda\hat{p}(\Lambda|H_0)\, d\Lambda \\ &= 1 - \hat{E}\{\Lambda|H_0\} + P_D'\,. \end{aligned}\tag{F.9}$$

where $\hat{E}\{\Lambda|H_0\}$ is the first moment of $\hat{p}(\Lambda|H_0)$.

The following example will illustrate the error incurred in using (F.6) and (F.8) instead of (F.6) and (F.7). We consider the problem of detecting a Gaussian signal in Gaussian noise in which the covariance matrices are diagonal and the samples have equal variances, as discussed by Van Trees (Van Trees, 1968). The likelihood ratio is

$$\Lambda(\underline{R}) = \frac{|K_0|^{1/2}}{|K_1|^{1/2}} \exp\left[\frac{1}{2\sigma_n^2}\frac{\sigma_s^2}{\sigma_n^2+\sigma_s^2}\sum_{i=1}^{M} R_i^2\right] \underset{H_0}{\overset{H_1}{\gtrless}} \eta \tag{F.10}$$

where the R_i are the received data, σ_n^2 and σ_s^2 are the noise and signal variances per sample, and K_0 and K_1 are covariance matricies. As a specific problem, we will investigate the case where $M = 2$ and $\sigma_s^2/\sigma_n^2 = 4$.

ROC's resulting from 25 Monte Carlo simulations of N = 2000 runs each were calculated. The output of each run was a single value of Λ conditioned to H_0, not just a sufficient statistic. The density $\hat{p}(\Lambda|H_0)$ was non-negative and its zeroth moment was constrained to be unity. No constraint was placed on the first moment. False alarm and detection probabilities were calculated as follows

$$\hat{P}_F = \frac{1}{N} \Sigma\, 1 \qquad (F.11)$$

$$\hat{P}'_D = \frac{1}{N} \Sigma\, \Lambda \qquad (F.12)$$

where the summations extend over all $\Lambda \geq \eta$, and

$$\hat{P}_D = 1 - \frac{1}{N} \Sigma\, \Lambda \qquad (F.13)$$

where the summation extends over all $\Lambda < \eta$. The mean value ROC's are shown in Figure F.1 along with two representative 95% confidence intervals on detection probability. The triangles represent mean $(\hat{P}'_D,\hat{P}_F)$ points calculated using (F.12) and (F.11); circles represent mean $(\hat{P}_D,\hat{P}_F)$ points calculated using (F.13) and (F.11). Note particularly the large variance associated with calculating detection probability via (F.12) as opposed to (F.13). A detailed account of $\hat{P}_F$, $\hat{P}'_D$, and $\hat{P}_D$ means and variances for the threshold values investigated is given in Table F.1. The mean $(\hat{P}_D,\hat{P}_F)$ points describe an ROC that is essentially equivalent to known analytic results (Van Trees, 1968) while the mean $(\hat{P}'_D,\hat{P}_F)$ points differ significantly.

Appendix

A theoretical model of a standard histogram estimator for $p(\Lambda|H_0)$ also suggests the use of (F.7) instead of (F.8) for the estimation of detection probability. As a specific example, the same Gaussian signal in Gaussian noise problems discussed in the text will be used.

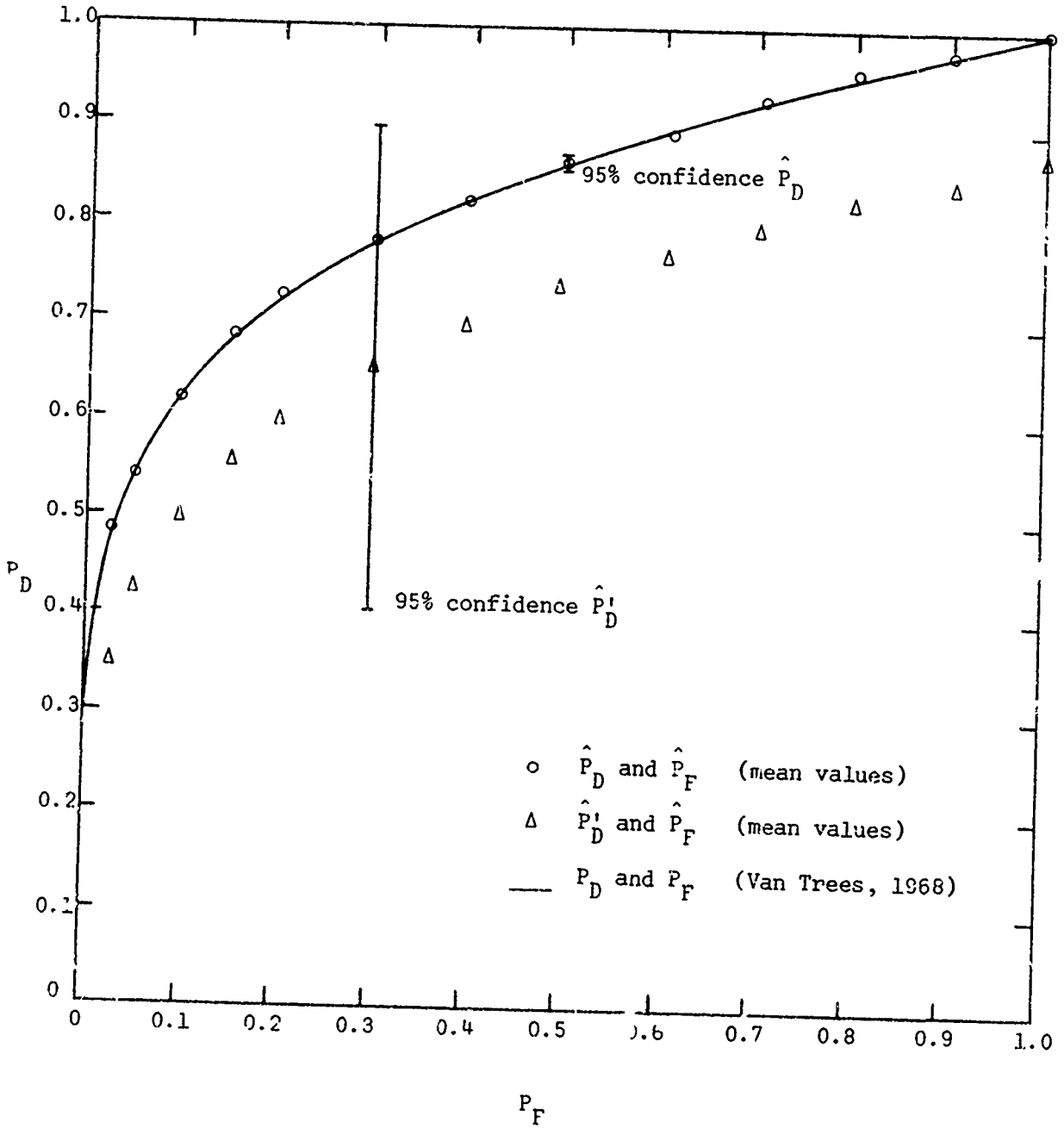


Figure F.1. Mean ROC's via Computer Simulation.

$\hat{P}_F$, $\hat{P}'_D$, and $\hat{P}_D$ Means and Variances

Threshold(η)	$\hat{P}_F$		$\hat{P}'_D$		$\hat{P}_D$	
	mean	var.	mean	var.	mean	var.
.200	1.000	.000 E-04	.868	.143 E-01	1.000	.000 E-05
.218	.900	.502	.947	.144	.979	.220
.239	.799	.717	.824	.144	.956	.333
.266	.697	1.420	.799	.145	.931	.780
.301	.597	1.490	.770	.146	.902	.857
.348	.495	1.500	.737	.144	.809	.941
.416	.395	.716	.699	.143	.831	.486
.524	.298	.615	.650	.141	.786	.832
.725	.200	.368	.594	.142	.726	1.030
.912	.150	.428	.554	.143	.686	2.346
1.262	.101	.292	.501	.143	.633	2.474
2.197	.050	.202	.419	.139	.551	7.652
3.825	.024	.085	.344	.141	.476	8.544

Table F.1

Let $\hat{p}(\Lambda|H_0)$ be our histogram estimator of $p(\Lambda|H_0)$ consisting of K = 250 cells each of width $\Delta\Lambda$ = .2 and an end cell representing all probability beyond the truncation point T = 50. The distribution of observations in the i^{th} cell is binomial with parameter p_i and, as a simplification, the distributions will be assumed independent from cell to cell. (In reality, the joint distribution of observations in all cells is a multinomial in the K + 1 parameter family of distributions.) The discrete versions of (F.6), (F.7), and (F.8) are

$$\hat{P}_F = \sum_{i=\eta'}^{K+1} P_i|H_0 \tag{F.14}$$

$$\hat{P}_D = 1 - \sum_{i=1}^{\eta'-1} \Lambda_i P_i|H_0 \tag{F.15}$$

$$\hat{P}'_D = \sum_{i=\eta'}^{K+1} \Lambda_i P_i|H_0 \tag{F.16}$$

where Λ_i and $P_i|H_0$ are the likelihood ratio and false alarm probability associated with the i^{th} cell and η' is the cell identifier corresponding to a threshold value η.

Our comparison of (F.15) and (F.16) will consist of calculating ROC's based on each and evaluating $\text{var}(\hat{P}_D)$ and $\text{var}(\hat{P}'_D)$ for various values of η'. The equation resulting in the larger value of bias and variance will be considered the least desirable of the two. For a single cell

$$\text{var}(P_i|H_0) = \frac{1}{N} p_i(1-p_i) \tag{F.17}$$

where N equals the total number of observations. Since independence between cells is assumed

$$\text{var}(\hat{P}_D) = \sum_{i=1}^{\eta'-1} \Lambda_i^2 \frac{1}{N} p_i(1-p_i) \tag{F.18}$$

$$\text{var}(\hat{P}'_D) = \sum_{i=\eta'}^{K+1} \Lambda_i^2 \frac{1}{N} p_i(1-p_i) . \tag{F.19}$$

Using a known analytic expression for $p(\Lambda|H_0)$ to derive values of p_i (Van Trees, 1968), the ROC's in Figure F.2 were calculated. Also shown are two representative 95% confidence intervals on detection probability. Note particularly the large variance associated with calculating detection probability via (F.16) as opposed to (F.15). A detailed account of $\hat{P}_F$, $\hat{P}'_D$, and $\hat{P}_D$ along with their variances is given in Table F.2. This analytic investigation complements the empirical results of the Monte Carlo simulation study. Both suggest that the calculation of detection proability with the aid of (F.3) is less biased and less variant when using (F.7) instead of (F.8) for data arising from a simulation of the likelihood ratio under H_0.

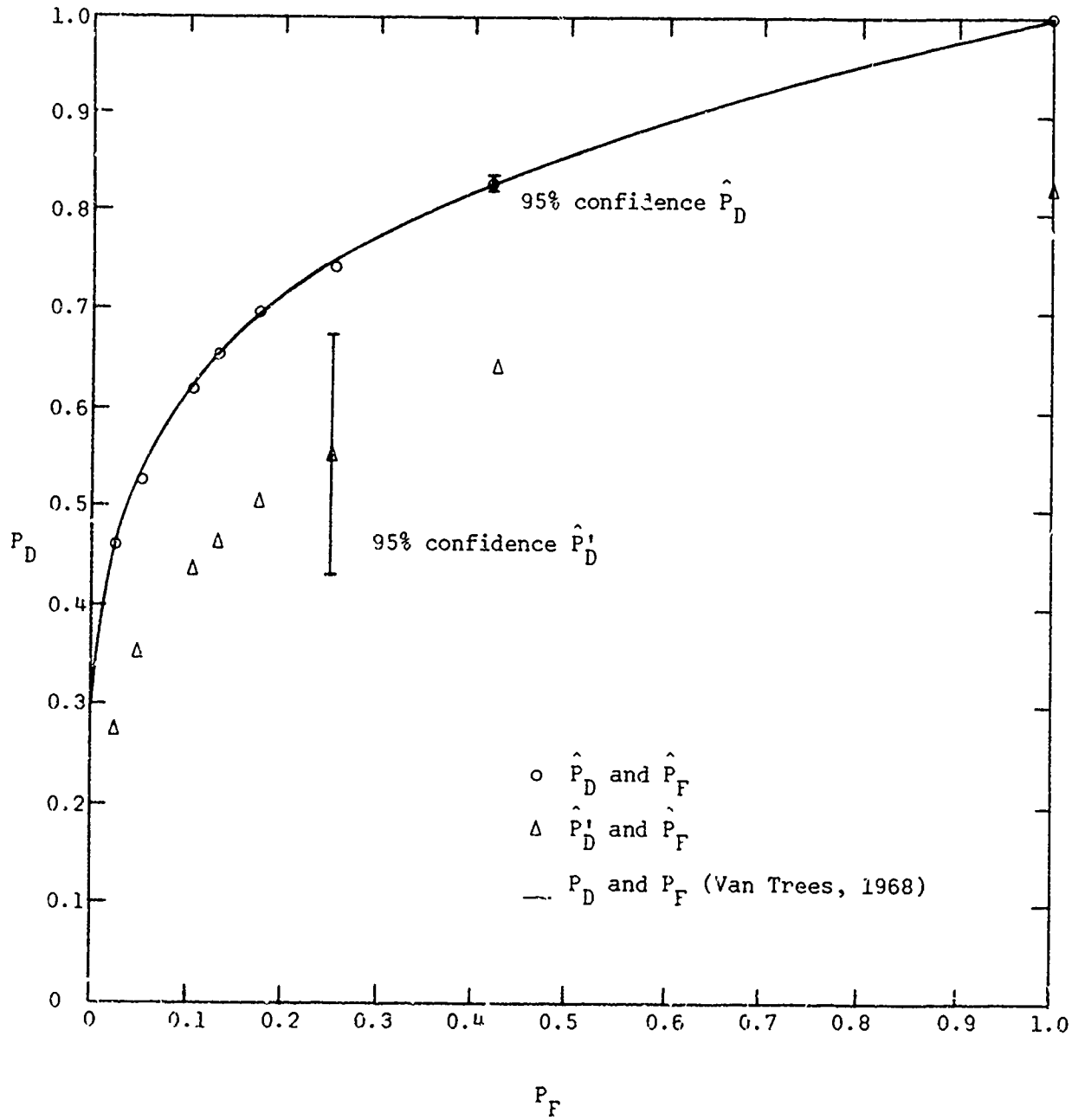


Figure F.2. ROC's from Histogram Model.

$\hat{P}_F$, $\hat{P}'_D$, and $\hat{P}_D$ and their Variances

Threshold(η')	$\hat{P}_F$	var.($\hat{P}_F$)	$\hat{P}'_D$	var.($\hat{P}'_D$)	$\hat{P}_D$	var.($\hat{P}_D$)
2	1.000	3.120 E-04	.818	.330 E-02	1.000	.000 E-04
3	.419	1.910	.644	.329	.826	.110
4	.252	1.210	.560	.327	.743	.284
5	.176	.862	.507	.326	.689	.457
6	.133	.656	.468	.324	.650	.624
7	.105	.523	.438	.322	.620	.784
12	.489	.244	.348	.315	.530	1.510
20	.242	.121	.278	.305	.460	2.530

Table F.2

Appendix G

COMPUTER PROGRAMS

The computer programs used in obtaining the optimal array processor performance reported in Chapter VIII are listed below. Four sections compromise this appenidx

(1) Gaussian Signal of Uncertain Direction in Gaussian Noise (GUD)

(2) Signal Known Except for Direction in Noise with an Additive Directional Noise Component of Known Direction (SKED in NKD)

(3) Signal Known Exactly in Noise with an Additive Directional Noise Component of Uncertain Direction (SKE in NUD).

The fourth section contains general programs which are shared by those in the first three plus the program which calculates detection and false alarm probabilities.

The Monte Carlo simulations were performed on a Digital Equipment Corporation PDP-9 computer. Each problem was broken into a sequence of two steps as illustrated in Figure 7.1. In the first step, Gaussian random numbers (mean zero, variance one) are read from a magnetic tape. For each simulation run, the collection $\{G_1/G_0, \underline{\delta}_1, \underline{\delta}_0\}$ is calculated and then written on another magnetic tape. Each collection is read during the second step and the remaining numerical integrations are performed. As they are calculated, the likelihood ratios are written on a DECtape which is eventually read by the program which calculates detection and false alarm probabilities. Several important concepts in the calculation of performance for likelihood ratio processors via computer simulation are discussed in Appendix F.

Gaussian Signal of Uncertain Direction in Gaussian Noise (GUD)

(See SKE in NUD Section for "SUBROUTINE GAUSS1")

```
C       GAUSSIAN SIGNAL OF UNCERTAIN DIRECTION IN GAUSSIAN NOI
C                                                 S
C       PART I
C
C       $A DTF1 1 ; $A DTF2 2
C       PROGRAM CALLS SUBROUTINES IN, OUT, PARAM1, ORDER, AND SKIP
C
C       INITIALIZING
C
        DIMENSION ZR(10), ZI(10), IND(100), ABR(10,10), ABI(10,10),
     1 G(10), ABRI(100), ABII(100)
        COMMON XI(250), X2(250)
        IXXX=251
        IYYY=1
        IBLOCK=0
        SQT=SQRT(2.)
C
C       INPUT PARAMETERS
C
        CALL PARAM1(ISKIP,M,K,IRUNS)
        IF(ISKIP.NE.0) CALL SKIP(1,IS(IP)
C
C       ORDER HANDLER FOR OUTPUTTING SUFFICIENT STATISTICS
C
        KI=K-1
        CALL ORDER(M,KI,IND,NUM)
C
        DO 200 JJ=1,IRUNS
        DO 140 M1=1,M
C
C       SINGLE FREQUENCY CONDITIONAL SOLUTION
C
        DO 100 I=1,K
        CALL IN(ZR(I),IXXX)
        IF(M1.EQ.1) ZR(I)=SQT*ZR(I)
        CALL IN(ZI(I),IXXX)
100     IF(M1.EQ.1) ZI(I)=0.
C
        G(M1)=0.
        DO 110 I=1,K
110     G(M1) = G(M1) + ZR(I)*ZR(I) + ZI(I)*ZI(I)
C
        DO 130 L=1,KI
        CLR=0.
        CLI=0.
        KL=K-L
        DO 120 I=1,KL
        IL=I+L
        CLR= CLR + ZR(I)*ZR(IL) + ZI(I)*ZI(IL)
120     CLI= CLI + ZR(I)*ZI(IL) - ZI(I)*ZR(IL)
        ABR(M1,L)=CLR
130     ABI(M1,L)=CLI
140     CONTINUE
```

```
C
C       COMBINE SINGLE FREQUENCY CONDITIONAL SOLUTIONS
C
        GG=0.
        DO 150 MM=1,M
150     GG=GG+G(JM)
        DO 160 I=1,100
        ABRI(I)=0.
160     ABII(I)=0.
        DO 180 MM=1,M
        DO 170 L=1,K1
        INDEX= (MM-1)*L + 1
        ABRI(INDEX)= ABRI(INDEX) + ABR(MM,L)
170     ABII(INDEX)= ABII(INDEX) + ABI(MM,L)
180     CONTINUE
C
C       OUTPUT
C
        CALL OUT(GG,IYYY,IBLOCK)
        DO 190 I=1,NUM
        J=IND(I)
        CALL OUT(ABRI(J),IYYY,IBLOCK)
190     CALL OUT(ABII(J),IYYY,IBLOCK)
C
C       END OF ONE SIMULATION RUN
C
        IF(JJ.NE.(JJ/100)*100) GO TO 200
        WRITE(6,195) JJ
195     FORMAT(2X,15HRUNS COMPLETED=,I5)
C
200     CONTINUE
C
        CALL OUT(0.,250,IBLOCK)
        WRITE(6,210) IBLOCK
210     FORMAT(/2X,7HIBLOCK=,I3)
        STOP
        END
```

```
C       SUBROUTINE INPUTS PARAMETERS FOR GAUSI
C
        SUBROUTINE PARAMI(ISKIP,M,K,IRUNS)
        WRITE(6,1)
1       FORMAT(/2X,36HBLOCKS (ISKIP) TO BE SKIPPED ON TAPE/)
        READ(5,5) ISKIP
5       FORMAT(I5)
        WRITE(6,7) ISKIP
7       FORMAT(/2X,6HISKIP=,I5)
        WRITE(6,10)
10      FORMAT(/2X,42HNUMBER OF FREQUENCIES (M) AND ELEMENTS (K)/)
        READ(5,20) M, K
20      FORMAT(2I2)
        WRITE(6,30) M, K
30      FORMAT(/2X,2HM=,I2,5X,2HK=,I2)
        WRITE(6,40)
40      FORMAT(/2X,14HNUMBER OF RUNS/)
        READ(5,50) IRUNS
50      FORMAT(I5)
        WRITE(6,60) IRUNS
60      FORMAT(/2X,5HRUNS=,I5)
        RETURN
        END
```

```
      GAUSSIAN SIGNAL OF UNCERTAIN DIRECTION IN GAUSSIAN NOISE
C       PART II
C
C       $A DTF1 1 } $A DTF2 2
C       PROGRAM CALLS SUBROUTINES IN, OUT, PARM2U, ORDER, SKIP,
C       BESNIS, GAUSS1, FG2U
C

C       INITIALIZING
C
        DIMENSION FI(50), PI(200), A(20,20)
        COMMON X1(250), X2(250), IND(100), ABR(100), ABI(100),
     1 C, AIOAO, CONST, AOCBO, AOSBO, M, NUM
        IXXX=251
        IYYY=1
        IBLOCK=0
C
C       INPUT PARAMETERS
C
        CALL PARM2U(ISKIP,M,K,IRUNS,S,AN,AO,BO)
        IF(ISKIP.NE.0) CALL SKIP(2,ISKIP)
C
C       ORDER HANDLER FOR INPUTTING SUFFICIENT STATISTICS
C
        K1=K-1
        CALL ORDER(M,K1,IND,NUM)
C
        C=S/(AN+FLOAT(K)*S)
        QO=AN**K
        Q1=(AN**(K-1))*(FLOAT(K)*S+AN)
        Q=(QO/Q1)**M
C
        DO 200 JJ=1,IRUNS
C
C       INPUT SUFFICIENT STATISTICS
C
        CALL IN(GG,IXXX)
        DO 10 I=1,NUM
        CALL IN(ABR(I),IXXX)
10      CALL IN(ABI(I),IXXX)
C
C       CALCULATE L
C
        IF(AO.LE.0.) AIOAO=1.
        IF(AO.LE.0.) GO TO 20
        CALL BESNIS(AO,2,FI,PI)
        AIOAO=FI(1)
20      CONST=FLOAT(K-1)/(2.*3.1415)*AIOAO)
        AOCBO=AO*COS(BO)
        AOSBO=AO*SIN(BO)
        YR041=3.14159/FLOAT(K-1)

        YR041=-YR041
```

```
      CALL GAUSS1(XRDM1,YRDM1,6,AREA)
C
      AL= Q * EXP((C/2.)*GG) * AREA
C
C     OUTPUT L
C
      CALL OUT(AL,IYYY,IBLOCK)
C
C     END OF ONE SIMULATION RUN
C
      IF(JJ.NE.(JJ/50)*50) GO TO 200
      WRITE(6,195) JJ
195   FORMAT(2X,15HRUNS COMPLETED=,I5)
C
200   CONTINUE
C
      CALL OUT(0.,250,IBLOCK)

      WRITE(6,210) IBLOCK
210   FORMAT(/2X,7HIBLOCK=,I3)
      STOP
      END
```

```
C       SUBROUTINE INPUTS PARAMETERS FOR GAUS2U
C
        SUBROUTINE PARM2U(ISKIP,M,K,IRUNS,S,AN,A0,B0)
        WRITE(6,1)
1       FORMAT(/2X,36HBLOCKS (ISKIP) TO BE SKIPPED ON TAPE/)
        READ(5,5) ISKIP
5       FORMAT(I5)
        WRITE(6,7) ISKIP
7       FORMAT(/2X,6HISKIP=,I5)
        WRITE(6,10)
10      FORMAT(/2X,42HNUMBER OF FREQUENCIES (M) AND ELEMENTS (K)/)
        READ(5,20) M,K
20      FORMAT(2I2)
        WRITE(6,30) M,K
30      FORMAT(/2X,2HM=,I2,5X,2HK=,I2)
        WRITE(6,40)
40      FORMAT(/2X,14HNUMBER OF RUNS/)
        READ(5,50) IRUNS
50      FORMAT(I5)
        WRITE(6,60) IRUNS
60      FORMAT(/2X,5HRUNS=,I5)
        WRITE(6,70)
70      FORMAT(/2X,34HPOWERS OF SIGNAL (S) AND NOISE (N)/)
        READ(5,80) S, AN
80      FORMAT(2F5.2)
        WRITE(6,90) S, AN
90      FORMAT(/2X,2HS=,F5.2,5X,2HN=,F5.2)
        WRITE(6,100)
100     FORMAT(/2X,32HUNCERTAINTY PARAMETERS A0 AND B0/)
        READ(5,110) A0, B0
110     FORMAT(2F5.2)
        WRITE(6,120) A0, B0
120     FORMAT(/2X,3HA0=,F5.2,5X,3HB0=,F5.2)
        RETURN
        END
```

```
C       FUNCTION CALLED BY GAUSS1 IN PROGRAM GAUS2U
C
        FUNCTION G1(WOTS)
        DIMENSION JN(10)
        COMMON X1(250), X2(250), IND(100),ABR(100), ABI(100),
     1 C, AIOAO, CONST, AOCBO, AOSBO, M, NUM
C
        SS=0.
C
C
        DO 10 I=1,NUM
        ARG= FLOAT(IND(I)-1) * WOTS
10      SS= SS + ABR(I)*COS(ARG) - ABI(I)*SIN(ARG)
        SS=C*SS + AOCBO*COS(FLOAT(M-1)*WOTS) - AOSBO*SIN(FLOAT(M-1)*WOTS)

        G1=CONST*EXP(SS)
        RETURN
        END
```

Signal Known Except for Direction in Noise with an Additive Directional Noise Component of Known Direction (SKED in NKD)

(See GUD Section for "FUNCTION G1(WCTS)")

```
C       SIGNAL KNOWN EXCEPT FOR DIRECTION
C       IN NOISE OF KNOWN DIRECTION
C       PART I
C
C       $A DTF1 1/DTF2 2
C       PROGRAM CALLS SUBROUTINES IN, OUT, PARA13, ORDER, AND SKIP
C
C
C       INITIALIZING
C
        REAL ILR, ILI, NWOTN
        DIMENSION ZR(10), ZI(10), IND(100), ABR(10,10), ABI(10,10),
     1 ABRI(100), ABII(100)
        COMMON X1(250), X2(250)
        IXXX=251
        IYYY=1
        IBLOCK=0
C
C       INPUT PARAMETERS
C
        CALL PARA13(ISKIP,M,K,IRUNS,ES,AN,D)
        IF(ISKIP.NE.0) CALL SKIP(1,ISKIP)
C
C       ORDER HANDLER FOR OUTPUTTING SUFFICIENT STATISTICS
C
        K1=K-1
        CALL ORDER(4,K1,IND,NUM)
C
        WRITE(6,10)
10      FORMAT(/2X,14HRUNS COMPLETED/)
        DO 200 JJ=1,IRUNS
C
C       A SINGLE FREQUENCY SIGNAL IS ASSUMED
C
C       SINGLE FREQUENCY CONDITIONAL SOLUTION
C
C       GENERATE FOURIER COEFFICIENTS UNDER H0
C
        NWOTN=3.14159/2.
        H0R=SQRT(ES/2.)
        H0I=0.
        CN=SQRT(AN/2.)
        CD=SQRT(D/2.)
        CALL IN(YDR,IYYY)
        CALL IN(YDI,IYYY)
        DO 100 I=1,K
        ARG=FLOAT(I-1)*NWOTN
        CS= COS(ARG)
        SN= SIN(ARG)
        CALL IN(YNR,IYYY)
        ZR(I)= CN*YNR + CD*(YDR*CS + YDI*SN)
        CALL IN(YNI,IYYY)
100     ZI(I)= CN*YNI + CD*(YDI*CS - YDR*SN)
```

```
C
C       SUFFICIENT STATISTICS INDEPENDENT OF TS
C
        C = D / (AN * (AN + FLOAT(K)*D))
        G1= -((FLOAT(K)*(AN + FLOAT(K-1)*D)) / (AN*(AN + FLOAT(K)*D)))
     1  * (BOR*BOR + BOI*BOI)
        G2= (2./AN) * (ZR(1)*BOR + ZI(1)*BOI)
        G3=0.
        DO 110 I=1,K
        ARG= FLOAT(I-1) * NWOTN
110     G3= G3 + (ZR(I)*BOR + ZI(I)*BOI)*COS(ARG) + (ZR(I)*BOI
     1  - ZI(I)*BOR)*SIN(ARG)
        G3= 2.*C*G3
        G= G1 + G2 - G3
C
C       SUFFICIENT STATISTICS DEPENDENT ON TS
C
        DO 130 L=1,K1
        L1= L+1
        GLR= (1./AN) * (ZR(L1)*BOR + ZI(L1)*BOI)
        GLI= (1./AN) * (ZR(L1)*BOI - ZI(L1)*BOR)
C
        HLR=0.
        HLI=0.
        DO 120 I=1,K
        ARG= FLOAT(I-L1) * NWOTN
        HLR= HLR + (ZR(I)*BOR + ZI(I)*BOI)*COS(ARG) + (ZR(I)*BOI
     1  - ZI(I)*BOR)*SIN(ARG)
120     HLI= HLI + (ZR(I)*BOI - ZI(I)*BOR)*COS(ARG)
     1  - (ZR(I)*BOR + ZI(I)*BOI)*SIN(ARG)
        HLR= C*HLR
        HLI= C*HLI
C
        KL=K-L
        ARG= FLOAT(L) * NWOTN
        CIL= C * FLOAT(KL) * (BOR*BOR + BOI*BOI)
        ILR= CIL * COS(ARG)
        ILI= CIL * SIN(ARG)
C
        ABR(1,L)= GLR - HLR + ILR
130     ABI(1,L)= -GLI + HLI - ILI
C
C       COMBINE SINGLE FREQUENCY SOLUTIONS
C
        DO 160 I=1,100
        ABRI(I)=0.
160     ABII(I)=0.
        DO 170 L=1,K1
        INDEX= (1-1)*L + 1
        ABRI(INDEX)= ABRI(INDEX) + ABR(1,L)
170     ABII(INDEX)= ABII(INDEX) - ABI(1,L)
C
C       OUTPUT
C
```

```
      CALL OUT(6,IYYY,IBLOCK)
      DO 190 I=1,NUM
      J=IND(I)
      CALL OUT(ABR1(J),IYYY,IBLOCK)
190   CALL OUT(ABI1(J),IYYY,IBLOCK)
C
C     END OF ONE SIMULATION RUN
C
      IF(JJ.NE.(JJ/100)*100) GO TO 200
      WRITE(6,195) JJ
195   FORMAT(2X,I5)
C
200   CONTINUE
      CALL OUT(0.,250,IBLOCK)
      WRITE(6,210) IBLOCK
210   FORMAT(/2X,7HIBLOCK=,I3)
      STOP
X     END
```

```
C       SUBROUTINE INPUTS PARAMETERS FOR SKED1
C
        SUBROUTINE PARAM3(ISKIP,M,K,IRUNS,ES,AN,D)
        WRITE(6,1)
1       FORMAT(/2X,36HBLOCKS (ISKIP) TO BE SKIPPED ON TAPE/)
        READ(5,5) ISKIP
5       FORMAT(I5)
        WRITE(6,7) ISKIP
7       FORMAT(/2X,6HISKIP=,I5)
        WRITE(6,10)
10      FORMAT(/2X,42HNUMBER OF FREQUENCIES (M) AND ELEMENTS (K)/)
        READ(5,20) M, K
20      FORMAT(2I2)
        WRITE(6,30) M, K
30      FORMAT(/2X,2HM=,I2,5X,2HK=,I2)
        WRITE(6,40)
40      FORMAT(/2X,14HNUMBER OF RUNS/)
        READ(5,50) IRUNS
50      FORMAT(I5)
        WRITE(6,60) IRUNS
60      FORMAT(/2X,6HRUNS=,I5)
        WRITE(6,70)
70      FORMAT(/2X,41HPOWERS OF SIGNAL (ES) AND NOISE (N AND D)/)
        READ(5,80) ES, AN, D
80      FORMAT(3F5.2)
        WRITE(6,90) ES, AN, D
90      FORMAT(/2X,3HES=,F5.2,5X,2HN=,F5.2,5X,2HD=,F5.2)
        RETURN
        END
```

```
C       SIGNAL KNOWN EXCEPT FOR DIRECTION
C       IN NOISE OF KNOWN DIRECTION
C       PART II
C
C       $A DTF1 1/DTF2 2
C       PROGRAM CALLS SUBROUTINES IN, OUT, PARA14, ORDER, SKIP,
C       BESNIS, R041, FG2U
C
C
C       INITIALIZING
C
        DIMENSION FI(50), PI(200), A(20,20)
        COMMON X1(250), X2(250), IND(100), ABR(100), ABI(100),
     1 C, AIOAO, CONST, AOCBO, AOSBO, M, NUM
        IXXX=251
        IYYY=1
        IBLOCK=0
C
C       INPUT PARAMETERS
C
        CALL PARA14(ISKIP,M,K,IRUNS,AO,BO)
        IF(ISKIP.NE.0) CALL SKIP(2,ISKIP)
C
C       ORDER HANDLER FOR INPUTTING SUFFICIENT STATISTICS
C
        K1=K-1
        CALL ORDER(M,K1,IND,NUM)
C
C
        C=2.
C
        WRITE(5,5)
5       FORMAT(/2X,14HRUNS COMPLETED/)
        DO 200 JJ=1,IRUNS
C
C       INPUT SUFFICIENT STATISTICS
C
        CALL IN(G,IXXX)
        DO 10 I=1,NUM
        CALL IN(ABR(I),IYYY)
10      CALL IN(ABI(I),IXXX)
C
C       CALCULATE L
C
C       UNCERTAIN DIRECTION
C
        IF(AO.LE.0.) AIOAO=1.
        IF(AO.LE.0.) GO TO 20
        CALL BESNIS(AO,2,FI,PI)
        AIOAO=FI(1)
20      CONST=FLOAT(M-1)/(2.*3.14159*AIOAO)
        AOCBO=AO*COS(BO)
        AOSBO=AO*SIN(BO)
        YR011=3.14159/FLOAT(M-1)
```

```
      XROM1=-YROM1
      CALL ROM1(XROM1,YROM1,A,I)
C
      AL= EXP(G) * A(I,I)
C
C     OUTPUT AL
C
      CALL OUT(AL,IYYY,IBLOCK)
C
C     END OF ONE SIMULATION RUN
C
      IF(JJ.NE.(JJ/50)*50) GO TO 200
      WRITE(6,195) JJ
195   FORMAT(2X,I5)
C
200   CONTINUE
C
      CALL OUT(0.,250,IBLOCK)
      WRITE(6,210) IBLOCK
210   FORMAT(/2X,7HIBLOCK=,I3)
      STOP
      END
```

```
C       SUBROUTINE INPUTS PARAMETERS FOR SKED2U
C
        SUBROUTINE PARAM(ISKIP,M,K,IRUNS,A0,B0)
        WRITE(6,1)
1       FORMAT(/2X,36HBLOCKS (ISKIP) TO BE SKIPPED ON TAPE/)
        READ(5,5) ISKIP
5       FORMAT(I5)
        WRITE(6,7) ISKIP
7       FORMAT(/2X,6HISKIP=,I5)
        WRITE(6,10)
10      FORMAT(/2X,42HNUMBER OF FREQUENCIES (M) AND ELEMENTS (K)/)
        READ(5,20) M,K
20      FORMAT(2I2)
        WRITE(6,30) M,K
30      FORMAT(/2X,2HM=,I2,5X,2HK=,I2)
        WRITE(6,40)
40      FORMAT(/2X,14HNUMBER OF RUNS/)
        READ(5,50) IRUNS
50      FORMAT(I5)
        WRITE(6,60) IRUNS
60      FORMAT(/2X,5HRUNS=,I5)
        WRITE(6,100)
100     FORMAT(/2X,32HUNCERTAINTY PARAMETERS A0 AND B0/)
        READ(5,110) A0, B0
110     FORMAT(2F5.2)
        WRITE(6,120) A0, B0
120     FORMAT(/2X,3HA0=,F5.2,5X,3HB0=,F5.2)
        RETURN
        END
```

```
      SUBROUTINE R041(X,Y,A,I)
C
C     THIS SUBROUTINE FORMS THE INTEGRAL OF G1(P) FROM X TO Y
C     THE ANSWER IS RETURNED IN A(I,I)
C     PROGRAM NEEDS FUNCTION G1(P)
C
      DIMENSION A(20,20)
      N=20
      A(1,1)=(G1(X)+G1(Y))*(Y-X)/2.
      DO 3 I=2,N
C
C     RECUR1
C
      SUM=0.
      IK=2**(I-1)-1
      DO 1 J=1,IK,2
      AJ=J
1     SUM=SUM+G1(X+((Y-X)/(2.**(I-1)))*AJ)
      A(I,1)=.5*(A(I-1,1)+((Y-X)/(2.**(I-2)))*SUM)
C
C     EXTRA1
C
      II=I-1
      DO 2 K=1,II
2     A(I,K+1)=((4.**K)*A(I,K)-A(I-1,K))/(4.**K-1.)
C
C
      IF(ABS(A(I,I)-A(I-1,I-1))-.1E-02) 5,3,3
3     CONTINUE
      WRITE(6,4)
4     FORMAT(//2X,46HWARNING * INTEGRAL 1 REQUIRED 20 ITERATIONS * )
5     RETURN
      END
```

Signal Known Exactly in Noise with an Additive Directional Noise Component of Uncertain Direction (SKE in NUD)

(See SKED in NKD Section for "SUBROUTINE PARAM4"; random numbers uniformly distributed between zero and one are read from DECtape by "SUBROUTINE IN3")

```
C       SIGNAL KNOWN EXACTLY IN NOISE WITH A DIRECTIONAL
C       NOISE COMPONENT OF UNCERTAIN DIRECTION

C

C       PART I    (STRUCTURED FOR A CHAIN)
C
C       $A DTF1 1/DTF2 2/DTF3 3
C       PROGRAM CALLS SUBROUTINES: IN,OUT,IN3,PARAMS,ORDER, AND SKIP
C
C       INITIALIZING
C
        REAL NWOTS, NWOTN
        COMMON X1(250), X2(250), X3(250), ZR(10), ZI(10), IND(100),
     1 ABR0(10,10), ABI0(10,10), ABR1(10,10), ABI1(10,10),
     2 G(10), ABR40(100), ABIM0(100), ABR41(100), ABI41(100)
        IXXX=251
        IYYY=1
        IBLOCK=0
        IXX3=251
        SQT=SQRT(2.)
C
C       INPUT PARAMETERS
C
        CALL PARAMS(ISKIP,M,K,IRUNS,ES,AN,D,,A0,B0)
        IF(ISKIP.NE.0)  CALL SKIP(1,ISKIP)
C
C       ORDER HANDLER FOR OUTPUTTING SUFFICIENT STATISTICS
C
        K1=K-1
        CALL ORDER(M,K1,IND,NDM)
C
        WRITE(6,10)
10      FORMAT(/2X,14HRUNS COMPLETED/)
        DO 200 JJ=1,IRUNS
C
C       CHOOSE A RANDOM NOISE DIRECTION
C
        CALL GOS(NWOTN,A0,IXX3)
        WOTN=NWOTN/FLOAT(M-1)
C
C       A SINGLE FREQUENCY SIGNAL IS ASSUMED AT (M-1)*W0=NW0
C
        NWOTS=3.14159/2.
        BOR=SQRT(ES/2.)
        BOI=0.
        CN=SQRT(AN/2.)
        CD=SQRT(D/2.)
        C1=(AN+FLOAT(K-1)*D)/(AN*(AN+FLOAT(K)*D))
        C2=D/(AN*(AN+FLOAT(K)*D))
C

        DO 181 MM=1,M
```

```
C
C       GENERATE FOURIER COEFFICIENTS UNDER HO
C
        CALL GEN(IX,IY,MM,WOTN,SQT,CN,CD)
C
C       SUFFICIENT STATISTICS INDEPENDENT OF TN
C
        CALL SUFITN(MM,C1,M,NWOTS,BOR,BOI,K,G1,G2)
C
C       SUFFICIENT STATISTICS DEPENDENT ON TN
C
        CALL SUFDTN(K1,K,MM,M,L,NWOTS,BOR,BOI,C2)
C
101     CONTINUE
C
C       COMBINE SINGLE FREQUENCY SOLUTIONS
C
        CALL COMB(G1,G2,G3,GG,M,K1)
C
C
        OUTPUT
C
        CALL OUTPUT(G3,GG,NUM,IYYY,IBLOCK)
C
C       END OF ONE SIMULATION RUN
C
        IF(JJ.NE.(JJ/100)*100) GO TO 200
        WRITE(6,195) JJ
195     FORMAT(2X,I5)
C
200     CONTINUE
        CALL OUT(0.,250,IBLOCK)
        WRITE(6,210) IBLOCK
210     FORMAT(/2X,7HIBLOCK=,I5)
        STOP
        END
```

```
C       SUBROUTINE INPUTS PARAMETERS FOR SKEDI
C
        SUBROUTINE PARAMS(ISKIP,M,K,IRUNS,ES,AN,D,A0,B0)
        WRITE(6,1)
1       FORMAT(/2X,36HBLOCKS (ISKIP) TO BE SKIPPED ON TAPE/)
        READ(5,5) ISKIP
5       FORMAT(I5)
        WRITE(6,7) ISKIP
7       FORMAT(/2X,6HISKIP=,I5)
        WRITE(6,10)
10      FORMAT(/2X,42HNUMBER OF FREQUENCIES (M) AND ELEMENTS (K)/)
        READ(5,20) M, K
20      FORMAT(2I2)
        WRITE(6,30) M, K
30      FORMAT(/2X,2HM=,I2,5X,2HK=,I2)
        WRITE(6,40)
40      FORMAT(/2X,14HNUMBER OF RUNS/)
        READ(5,50) IRUNS
50      FORMAT(I5)
        WRITE(6,60) IRUNS
60      FORMAT(/2X,5HRUNS=,I5)
        WRITE(6,70)
70      FORMAT(/2X,41HPOWERS OF SIGNAL (ES) AND NOISE (N AND D)/)
        READ(5,80) ES, AN, D
80      FORMAT(3F5.2)
        WRITE(6,90) ES, AN, D
90      FORMAT(/2X,3HES=,F5.2,5X,2HN=,F5.2,5X,2HD=,F5.2)
        WRITE(6,100)
100     FORMAT(/2X,32HUNCERTAINTY PARAMETERS A0 AND B0/)
        READ(5,110) A0, B0
110     FORMAT(2F5.2)
        WRITE(6,120) A0, B0
120     FORMAT(/2X,3HA0=,F5.2,5X,3HB0=,F5.2)
        RETURN
        END
```

```
C       SUBROUTINE CHOOSES A RANDOM NOISE DIRECTION
C
C       SUBROUTINE CALLS SUBROUTINE IN3
C       $A DTF3 3
C
        SUBROUTINE POS(NWOTN,A0,IXX3)
        REAL NWOTN
        DIMENSION A(25), B(25)
        COMMON X1(250),X2(250),X3(250)
C
        A(1)=.0096
        A(2)=.0290
        A(3)=.0485
        A(4)=.0680
        A(5)=.0877
        A(6)=.1077
        A(7)=.1280
        A(8)=.1486
        A(9)=.1698
        A(10)=.1914
        A(11)=.2136
        A(12)=.2368
        A(13)=.2607
        A(14)=.2858
        A(15)=.3121
        A(16)=.3400
        A(17)=.3700
        A(18)=.4023
        A(19)=.4379
        A(20)=.4775
        A(21)=.5230
        A(22)=.5772
        A(23)=.6458
        A(24)=.7431
        A(25)=.9317
C
        B(1)=.0153
        B(2)=.0459
        B(3)=.0766
        B(4)=.1075
        B(5)=.1387
        B(6)=.1703
        B(7)=.2025
        B(8)=.2353
        B(9)=.2689
        B(10)=.3034
        B(11)=.3391
        B(12)=.3760
        B(13)=.4144
        B(14)=.4548
        B(15)=.4975
        B(16)=.5429
        B(17)=.5918
```

```
      B(18)=.6450
      B(19)=.7038
      B(20)=.7704
      B(21)=.8475
      B(22)=.9411
      B(23)=1.0629
      B(24)=1.2425
      B(25)=1.6350
C
      CALL IN3(Y,IXX3)
      J=INT(50.*Y)+1
      IF(J.GT.50)  J=50
      IF(A0.EQ.0.) NWOTN=6.28318*(Y-.5)
      IF(A0.EQ.0.) GO TO 10
      IF(J.GT.25) GO TO 5
      IF(A0.EQ.7.) NWOTN=A(J)
      IF(A0.EQ.3.) NWOTN=B(J)
      GO TO 10
5     J=J-25
      IF(A0.EQ.7.) NWOTN=-A(J)
      IF(A0.EQ.3.) NWOTN=-B(J)
10    RETURN
      END
```

```
C       GENERAL INPUT HANDLER
C
C       SET IXX3=251 TO INITIALIZE
C       COMMON X1(250),X2(250),X3(250)
C       $A DTF3 3 ; READS FROM DT3
C
        SUBROUTINE IN3(Y,IXX3)
        COMMON X1(250), X2(250), X3(250)
        IF(IXX3.NE.251) GO TO 10
        READ(3) (X3(J), J=1,250)
         XX3=1
10      Y=X3(IXX3)
        IXX3=IXX3+1
        RETURN
        END
```

```
      SUBROUTINE GEN(IXXX,K,MM,WOTN,SQT,CN,CD)
C
      COMMON X1(250),X2(250),X3(250), ZR(10), ZI(10), IND(100),
     1 ABR0(10,10), ABI0(10,10), ABR1(10,10), ABI1(10,10),
     2 G(10), ABRM0(100), ABIM0(100), ABRM1(100), ABIM1(100)
C
C     GENERATE FOURIER COEFFICIENTS UNDER H0
C
      CALL IN(YDR,IXXX)
      CALL IN(YDI,IXXX)
      DO 100 I=1,K
      ARG=FLOAT(I-1)*FLOAT(MM-1)*WOTN
      CS=COS(ARG)
      SN=SIN(ARG)
      CALL IN(YNR,IXXX)
      ZR(I)=CN*YNR + CD*(YDR*CS + YDI*SN)
      IF(MM.EQ.1) ZR(I)=SQT*ZR(I)
      CALL IN(YNI,IXXX)
      ZI(I)=CN*YNI + CD*(YDI*CS - YDR*SN)
100   IF(MM.EQ.1) ZI(I)=0.
C
      RETURN
      END
```

```
      SUBROUTINE SUFITN(MM,C1,M,NWOTS,BOR,BOI,K,G1,G2)
C
      REAL NWOTS
      COMMON X1(250), X2(250), X3(250), ZR(10), ZI(10), IND(100),
     1 ABRO(10,10), ABIO(10,10), ABR1(10,10), ABI1(10,10),
     2 G(10), ABR40(100), ABI40(100), ABR41(100), ABI41(100)
C
C     SUFFICIENT STATISTICS INDEPENDENT OF TN
C
      G(MM)=0.
      DO 110 I=1,K
110   G(MM)=G(MM) + ZR(I)*ZR(I) + ZI(I)*ZI(I)
      G(MM)= -C1*G(44)
      IF(M.NE.4) GO TO 130
      G1=0.
      DO 120 I=1,K
      ARG=FLOAT(I-1)*NWOTS
120   G1=G1 + (ZR(I)*BOR + ZI(I)*BOI) * COS(ARG)
     1      + (ZR(I)*BOI - ZI(I)*BOR) * SIN(ARG)
      G1=2.*C1*G1
      G2= -FLOAT(K)*C1*(BOR*BOR + BOI*BOI)
130   CONTINUE
C
      RETURN
      END
```

```
      SUBROUTINE SUFDTN(K1,K,MM,M,L,NWOTS,BOR,BOI,C2)
C
      REAL NWOTS
      COMMON X1(250), X2(250), X3(250), ZR(10), ZI(10), IND(100),
     1 ABRO(10,10), ABIO(10,10), ABR1(10,10), ABI1(10,10),
     2 G(10), ABRM0(100), ABIM0(100), ABRM1(100), ABIM1(100)
C
C     SUFFICIENT STATISTICS DEPENDENT ON TN
C
      DO 180 L=1,K1
C
      CLR=0.
      CLI=0.
      KL=K-L
      DO 140 I=1,KL
      IL=I+L
      CLR=CLR + ZR(I)*ZR(IL) + ZI(I)*ZI(IL)
140   CLI=CLI + ZR(I)*ZI(IL) - ZI(I)*ZR(IL)
C
      DLR=0.
      DLI=0.
      ELR=0.
      ELI=0.
      FLR=0.
      FLI=0.
      IF(MM.NE.M) GO TO 170
C
      LIM=L-1
      DO 150 I=1,KL
      ARG=FLOAT(I+LIM)*NWOTS
      DLR=DLR + (ZR(I)*BOR + ZI(I)*BOI) * COS(ARG)
     1        + (ZR(I)*BOI - ZI(I)*BOR) * SIN(ARG)
150   DLI=DLI + (ZR(I)*BOI - ZI(I)*BOR) * COS(ARG)
     1        - (ZR(I)*BOR + ZI(I)*BOI) * SIN(ARG)
C
      LIP=L+1
      DO 160 I=LIP,K
      ARG=FLOAT(I-LIP)*NWOTS
      ELR=ELR + (ZR(I)*BOR + ZI(I)*BOI) * COS(ARG)
     1        + (ZR(I)*BOI - ZI(I)*BOR) * SIN(ARG)
160   ELI=ELI + (ZR(I)*BOI - ZI(I)*BOR) * COS(ARG)
     1        - (ZR(I)*BOR + ZI(I)*BOI) * SIN(ARG)
C
      ARG=FLOAT(L)*NWOTS
      CFL=FLOAT(KL)*(BOR*BOR + BOI*BOI)
      FLR=CFL*COS(ARG)
      FLI= -CFL*SIN(ARG)
C
170   CONTINUE
      ABR1(MM,L)= 2.*C2*(CLR-DLR-ELR+FLR)
      ABI1(MM,L)= 2.*C2*(CLI-DLI+ELI+FLI)
C
```

```
      A480(11,L)= 2.*C2*CL9
180   AB10(11,L)= 2.*C2*CL1
C
      RETURN
      END
```

```
      SUBROUTINE COMB(G1,G2,G3,GG,M,KI)
C
      COMMON X1(250), X2(250), X3(250), ZR(10), ZI(10), IND(100),
     1 ABRO(10,10), ABIO(10,10), ABRI(10,10), ABII(10,10),
     2 G(10), ABR40(100), ABI40(100), ABR41(100), ABI41(100)
C
C     COMBINE SINGLE FREQUENCY SOLUTIONS
C
      DO 182 I=1,100
      ABR41(I)=0.
      ABI41(I)=0.
      ABR40(I)=0.
182   ABI40(I)=0.
C
      G3=0.
      DO 183 MM=1,M
183   G3=G3 + G(MM)
      GG=G1 + G2
      DO 185 MM=1,M
      DO 184 L=1,KI
      INDEX= (MM-1)*L + I
      ABR41(INDEX)=ABR41(INDEX) + ABRI(MM,L)
      ABI41(INDEX)=ABI41(INDEX) + ABII(MM,L)
      ABR40(INDEX)=ABR40(INDEX) + ABRO(MM,L)
184   ABI40(INDEX)=ABI40(INDEX) + ABIO(MM,L)
185   CONTINUE
C
      RETURN
```

```
      SUBROUTINE OUTPUT(G3,GG,NUM,IYYY,IBLOCK)
C
      COMMON X1(250), X2(250), X3(250), Z(10), ZI(10), IND(100),
     1 ABR0(10,10), ABI0(10,10), ABR1(10,10), ABI1(10,10),
     2 G(10), ABRM0(100), ABIM0(100), ABRM1(100), ABIM1(100)
C
C     OUTPUT
C
      CALL OUT(G3,IYYY,IBLOCK)
      CALL OUT(GG,IYYY,IBLOCK)
      DO 190 I=1,NUM
      J=IND(I)
      CALL OUT(ABRM1(J),IYYY,IBLOCK)
      CALL OUT(ABIM1(J),IYYY,IBLOCK)
      CALL OUT(ABRM0(J),IYYY,IBLOCK)
190   CALL OUT(ABIM0(J),IYYY,IBLOCK)
C
      RETURN
```

```
C       SIGNAL KNOWN EXACTLY IN NOISE WITH A DIRECTIONAL
C       NOISE COMPONENT OF UNCERTAIN DIRECTION


C
C       PART II
C
C       $A DTF: I/DTF2 2
C        PROGRAM CALLS SUBROUTINES: IN,OUT,PARA44,ORDER,SKIP,BESNIS,
C        GAUSS1, GAUSS2, G1, G2
C

CC       INITIALIZING
C
        DIMENSION FI(50), PI(200), A(20,20), X(4), Y(4)
        COMMON X1(250), X2(250), IND(100), ABR10(100), ABI10(100),
     1 ABR11(100), ABI11(100), C, AIOA0, CONST, A0CB0, A0SB0, M,NUM
        IXXX=251
        IYYY=1
        IBLOCK=0
C
C       INPUT PARAMETERS
C
        CALL PARA44(ISKIP,M,K,IRUNS,A0,B0)
        IF(ISKIP.NE.0) CALL SKIP(2,ISKIP)
C
C       ORDER HANDLER FOR INPUTTING SUFFICIENT STATISTICS
C
        K1=K-1
        CALL ORDER(M,K1,IND,NUM)
C
        C=1.
        IF(A0.LE.0.) AIOA0=1.
        IF(A0.LE.0.) GO TO 20
        CALL BESNIS(A0,2,FI,PI)
        AIOA0=FI(1)
20      CONST=FLOAT(M-1)/(2.*3.14159*AIOA0)
        A0CB0=A0*COS(B0)
        A0SB0=A0*SIN(B0)
        YR041=3.14159/FLOAT(M-1)
        XR041=-YR041
        X(1)=XR041
        X(2)=XR041/2.
        X(3)=0.
        X(4)=YR0M1/2.
        Y(1)=X(2)
        Y(2)=X(3)
        Y(3)=X(4)
        Y(4)=YR0M1
C
        WRITE(6,5)
5       FORMAT(/2X,14HRUNS COMPLETED/)
        DO 200 JJ=1,IRUNS
```

```
C
C       INPUT SUFFICIENT STATISTICS
C
        CALL IN(G3,IXXX)
        CALL IN(GG,IXXX)
        DO 10 I=1,NUM
        CALL IN(ABRM1(I),IXXX)
        CALL IN(ABIM1(I),IXXX)
        CALL IN(ABRM0(I),IXXX)
10      CALL IN(ABIM0(I),IXXX)
C
C       CALCULATE L
C
        ANUM=0.
        ADENOM=0.
        DO 15 I=1,4
        CALL GAUSS1(X(I),Y(I),6,AREA)
15      ANUM=ANUM+AREA
        DO 17 I=1,4
        CALL GAUSS2(X(I),Y(I),6,AREA)
17      ADENOM=ADENOM+AREA
C
        AL= EXP(GG) * ANUM/ADENOM
C
C       OUTPUT L
C
        CALL OUT(AL,IYYY,IBLOCK)
C
C       END OF ONE SIMULATION RUN
C
        IF(JJ.NE.(JJ/50)*50) GO TO 200
        WRITE(6,195) JJ
195     FORMAT(2X,I5)
C
200     CONTINUE
C
        CALL OUT(0.,250,IBLOCK)
        WRITE(6,210) IBLOCK
210     FORMAT(/2X,7HIBLOCK=,I3)
        STOP
        END
```

```
C       SUBROUTINE GAUSS1(X,Y,M,AREA)
C       SUBROUTINE GAUSS USES THE M-POINT GAUSS-LEGENDRE QUADRATURE
C       FORMULA TO COMPUTE THE INTEGRAL OF G1(X)*DX BETWEEN
C       INTEGRATION LIMITS X AND Y.  THE ROOTS OF SEVEN LEGEND
E                                    R
C       POLYNOMIALS AND THE WEIGHT FACTORS FOR THE CORRESPONDING
C       QUADRATURES ARE STORED IN THE Z AND WEIGHT ARRAYS
C       RESPECTIVELY.  M MAY ASSUME VALUES 2,3,4,5,6,10, AND 15
C       ONLY.  THE APPROPRIATE VALUES FOR THE M-POINT FORMULA ARE
C       LOCATED IN ELEMENTS Z(KEY(I), ... , Z(KEY(I+1)-1) AND
C       WEIGHT(KEY(I)), ... , WEIGHT(KEY(I+1)-1) WHERE THE PROPER
C       VALUE OF I IS DETERMINED BY FINDING THE SUBSCRIPT OF THE
C       ELEMENT OF THE ARRAY NPOINT WHICH HAS THE VALUE M.  IF AN
C       INVALID VALUE OF M IS USED, A TRUE ZERO IS RETURNED AS THE
C       VALUE OF GAUSS.
C
        DIMENSION NPOINT(7), KEY(8), Z(24), WEIGHT(24)
C
C       PRESET NPOINT, KEY, Z, AND WEIGHT ARRAYS
C
        DATA KEY(1),KEY(2),KEY(3),KEY(4),KEY(5),KEY(6),KEY(7),KEY(8)
     1           / 1, 2, 4, 6, 9, 12, 17, 25 /
C
        DATA NPOINT(1),NPOINT(2),NPOINT(3),NPOINT(4),NPOINT(5)
                                           ,
     1 NPOINT(6),NPOINT(7)
     2              / 2, 3, 4, 5, 6, 10, 15 /
C
        DATA Z(1), Z(2), Z(3), Z(4), Z(5), Z(6), Z(7), Z(8), Z(9),Z(10),
     1     Z(11),Z(12),Z(13),Z(14),Z(15),Z(16),Z(17),Z(18),Z(19),Z(20),
     2     Z(21),Z(22),Z(23),Z(24)
     3              / 0.577350, 0.0      , 0.774597,
     4        0.339981, 0.861136, 0.0      , 0.538469,
     5        0.906180, 0.238619, 0.661209, 0.932470,
     6        0.148874, 0.433395, 0.679410, 0.865063,
     7        0.973907, 0.0      , 0.201194, 0.394151,
     8        0.570972, 0.724418, 0.848207, 0.937273,
     6        0.987993 /
C
        DATA WEIGHT(1), WEIGHT(2), WEIGHT(3), WEIGHT(4), WEIGH
(5),
                                  T
     1      WEIGHT(6), WEIGHT(7), WEIGHT(8), WEIGHT(9),WEIGHT
10),
                                  (
     2      WEIGHT(11),WEIGHT(12),WEIGHT(13),WEIGHT(14),WEIGH
(15),
                                  T
     3      WEIGHT(16),WEIGHT(17),WEIGHT(18),WEIGHT(19),WEIGH
(20),
                                  T
     4      WEIGHT(21),WEIGHT(22),WEIGHT(23),WEIGHT(24)
     5              / 1.0      , 0.888889, 0.555556,
     6        0.652145, 0.347855, 0.568889, 0.478629,
     7        0.236927, 0.467914, 0.360762, 0.171324,
```

```
     8         0.295524, 0.269267, 0.219086, 0.149451,
     9         0.066671, 0.202574, 0.198431, 0.186161,
     A          0.166269, 0.139571, 0.107159, 0.070366,
     B         0.030753 /
C
C      FIND SUBSCRIPT OF  FIRST Z AND WEIGHT VALUE
C
       DO 1 I=1,7
       IF(M.EQ.NPOINT(I)) GO TO 2
1      CONTINUE
C
C      INVALID M USED
C
       AREA=0.
       RETURN
C
C      SET UP INITIAL PARAMETERS
C
2      JFIRST=KEY(I)
       JLAST=KEY(I+1) - 1
       C=(Y-X)/2.
       D=(Y+X)/2.
C
C      ACCUMULATE THE SUM IN THE M-POINT FORMULA
C
       SUM=0.
       DO 5 J=JFIRST,JLAST
       IF(Z(J).EQ.0.) SUM=SUM + WEIGHT(J)*G1(D)
5      IF(Z(J).NE.0.) SUM=SUM + WEIGHT(J)*(G1(Z(J)*C + D)
     1 + G1(-Z(J)*C + D))
C
C      MAKE INTERVAL CORRECTION AND RETURN
C
       AREA=C*SUM
       RETURN
       END
```

```
C       FUNCTION CALLED BY  GAUSS1  (NUMERATOR)
C
        FUNCTION G1(WOTS)
        COMMON X1(250), X2(250), IND(100), ABRM0(100), ABIM0(100),
     1 ABRM1(100), ABIM1(100), C, A0A0, CONST, A0CB0, A0SB0, M, NU1
C
        SS=0.
        DO 10 I=1,NUM
        ARG= FLOAT(IND(I)-1) * WOTS
10      SS= SS + ABRM1(I)*COS(ARG) - ABIM1(I)*SIN(ARG)
        SS=C*SS + A0CB0*COS(FLOAT(M-1)*WOTS) - A0SB0*SIN(FLOAT(M-1)*WOTS)

        G1=CONST*EXP(SS)
        RETURN
        END
```

```
      SUBROUTINE GAUSS2(X,Y,M,AREA)
C
C     SUBROUTINE GAUSS USES THE M-POINT GAUSS-LEGENDRE QUADRATURE
C     FORMULA TO COMPUTE THE INTEGRAL OF G2(X)*DX BETWEEN
C     INTEGRATION LIMITS X AND Y.  THE ROOTS OF SEVEN LEGEND
E                              R
C     POLYNOMIALS AND THE WEIGHT FACTORS FOR THE CORRESPONDING
C     QUADRATURES ARE STORED IN THE Z AND WEIGHT ARRAYS
C     RESPECTIVELY.  M MAY ASSUME VALUES 2,3,4,5,6,10, AND 15
C     ONLY.  THE APPROPRIATE VALUES FOR THE M-POINT FORMULA ARE
C     LOCATED IN ELEMENTS Z(KEY(I), ... , Z(KEY(I+1)-1) AND
C     WEIGHT(KEY(I)), ... , WEIGHT(KEY(I+1)-1) WHERE THE PROPER
C     VALUE OF I IS DETERMINED BY FINDING THE SUBSCRIPT OF THE
C     ELEMENT OF THE ARRAY NPOINT WHICH HAS THE VALUE M.  IF AN
C     INVALID VALUE OF M IS USED, A TRUE ZERO IS RETURNED AS THE
C     VALUE OF GAUSS.
C
      DIMENSION NPOINT(7), KEY(8), Z(24), WEIGHT(24)
C
C     PRESET NPOINT, KEY, Z, AND WEIGHT ARRAYS
C
      DATA KEY(1),KEY(2),KEY(3),KEY(4),KEY(5),KEY(6),KEY(7),KEY(8)
     1          / 1, 2, 4, 6, 9, 12, 17, 25 /
C
      DATA NPOINT(1),NPOINT(2),NPOINT(3),NPOINT(4),NPOINT(5),
     1 NPOINT(6),NPOINT(7)
     2             / 2, 3, 4, 5, 6, 10, 15 /
C
      DATA Z(1), Z(2), Z(3), Z(4), Z(5), Z(6), Z(7), Z(8), Z(9),Z(10),

     1     Z(11),Z(12),Z(13),Z(14),Z(15),Z(16),Z(17),Z(18),Z(19),Z(20),

     2     Z(21),Z(22),Z(23),Z(24)
     3              / 0.577350, 0.0     , 0.774597,
     4        0.339981, 0.861136, 0.0     , 0.538469,
     5        0.906180, 0.238619, 0.661209, 0.932470,
     6        0.148874, 0.433395, 0.679410, 0.865063,
     7        0.973907, 0.0     , 0.201194, 0.394151,
     8        0.570972, 0.724418, 0.348207, 0.937273,
     6        0.987993 /
C
      DATA WEIGHT(1), WEIGHT(2), WEIGHT(3), WEIGHT(4), WEIGH
(5),
                              T
     1      WEIGHT(6), WEIGHT(7), WEIGHT(8), WEIGHT(9),WEIGHT
10),
                              (
     2      WEIGHT(11),WEIGHT(12),WEIGHT(13),WEIGHT(14),WEIGH
(15),
                             T
     3      WEIGHT(16),WEIGHT(17),WEIGHT(18),WEIGHT(19),WEIGH
(20),
                             T
     4      WEIGHT(21),WEIGHT(22),WEIGHT(23),WEIGHT(24)
     5              / 1.0     , 0.888889, 0.555556,
     6        0.652145, 0.347855, 0.568889, 0.478629,
     7        0.236927, 0.467914, 0.360762, 0.171324,
```

```
     3         0.295524, 0.269267, 0.219086, 0.149451,
     9         0.066671, 0.202578, 0.198431, 0.186161,
     A          0.166269, 0.139571, 0.107159, 0.070366,
     B         0.030753 /
C
C       FIND SUBSCRIPT OF FIRST Z AND WEIGHT VALUE
C
        DO 1 I=1,7
        IF(M.EQ.NPOINT(I)) GO TO 2
1       CONTINUE
C
C       INVALID M USED
C
        AREA=0.
        RETURN
C
C       SET UP INITIAL PARAMETERS
C
2       JFIRST=KEY(I)
        JLAST=KEY(I+1) - 1
        C=(Y-X)/2.
        D=(Y+X)/2.
C
C       ACCUMULATE THE SUM IN THE M-POINT FORMULA
C
        SUM=0.
        DO 5 J=JFIRST,JLAST
        IF(Z(J).EQ.0.) SUM=SUM + WEIGHT(J)*G2(D)
5       IF(Z(J).NE.0.) SUM=SUM + WEIGHT(J)*(G2(Z(J)*C + D)
     1 + G2(-Z(J)*C + D))
C
C       MAKE INTERVAL CORRECTION AND RETURN
C
        AREA=C*SUM
        RETURN
        END
```

```
C       FUNCTION CALLED BY GAUSS2  (DENOMINATOR)

        FUNCTION G2(WOTS)
        COMMON X1(250), X2(250), IND(100), ABR10(100), ABI10(100),
     1  ABR11(100), ABI11(100), C, AIOAO, CONST, AOCBO, AOSBO, M, NUM
C
        SS=0.
        DO 10 I=1,NUM
        ARG= FLOAT(IND(I)-1) * WOTS
10      SS= SS + ABR10(I)*COS(ARG) - ABI10(I)*SIN(ARG)
        SS=C*SS + AOCBO*COS(FLOAT(M-1)*WOTS) - AOSBO*SIN(FLOAT(M-1)*WOTS)

        G2=CONST*EXP(SS)
        RETURN
        END
```

Supporting Programs

```
C       GENERAL INPUT HANDLER
C
C       SET IXXX=251 TO INITIALIZE
C       COMMON X(250)
C       $A DTF2 2 ; READS FROM DT2
C

        SUBROUTINE IN(Y,IXXX)
        COMMON X(250)
        IF(IXXX.NE.251) GO TO 10
        READ(2) (X(J), J=1,250)
        IXXX=1
10      Y=X(IXXX)
        IXXX=IXXX+1
        RETURN
        END
```

```
C       GENERAL OUTPUT HANDLER
C
C       SET IYYY=1, IBLOCK=0 TO INITIALIZE
C       COMMON X1(250), X(250)
C       SA DTF1 1 ) WRITES ON DT1
C
        SUBROUTINE OUT(Y,IYYY,IBLOCK)
        COMMON X1(250), X(250)
        X(IYYY)=Y
        IYYY=IYYY+1
        IF(IYYY.NE.251) GO TO 10
        WRITE(1) (X(J), J=1,250)
        IYYY=1
        IBLOCK=IBLOCK+1
10      RETURN
        END
```

```
C       ORDER HANDLER
C
C       WRITES IN INCREASING ORDER THE L*N+1 PRODUCT VALUES
C
        SUBROUTINE ORDER(M,K1,IN,NUM)
        DIMENSION IN(100)
        NUM=0

        DO 10 I=1,100
10      IN(I)=0
        DO 30 M1=1,M
        DO 20 L=1,K1
        IND= (M1-1)*L+1
20      IN(IND)=1
30      CONTINUE
        DO 40 I=1,100
        IF(IN(I).NE.1) GO TO 40
        NUM=NUM+1
        IN(NUM)=I
40      CONTINUE
        WRITE(6,50)
50      FORMAT(//2X,20HL*N+1 PRODUCT VALUES/)
        WRITE(6,60) (IN(J), J=1,NUM)
60      FORMAT(2X,5(I3,5X))
        RETURN
        END
```

```
C       SUBROUTINE TO SKIP ISKIP BLOCKS OF
C       250 NUMBERS/BLOCK STORED ON DTI
C
C       $A DTFI I
        SUBROUTINE SKIP(I,ISKIP)
        DIMENSION X(250)
        DO 10 K=1,ISKIP
10      READ(I) (X(J), J=1,250)
        RETURN
        END
```

```
      SUBROUTINE BESNIS(X, NAX,FI,PI)
C
C     THIS SUBROUTINE CALCULATES THE MODIFIED BESSEL FUNCTION
C     FI(1)=I0(X), FI(2)=I1(X), ..., FI(NMAX)=I(NMAX-1)(X)
C     DIVIDE BY 0 ERROR IF X=0.
C     DOES NOT CHECK FOR X LESS THAN 0.
C
      DIMENSION FI(50), PI(200)
      SUM=0.
      I=X
      JMAX=I+21
      TZ=2./X
      PI(JMAX+2)=0.
      PI(JMAX+1)=1.E-20
      DO 1 J=1,JMAX
      K=JMAX+2-J
      DK=K-1
      PI(K-1)=DK*TZ*PI(K)+PI(K+1)
1     SUM=SUM+PI(K)
      SUM=SUM+SUM
      A=EXP(X)/(PI(1)+SUM)
      DO 2 N=1,NMAX
2     FI(N)=A*PI(N)
      RETURN
      END
```

```
C       PROGRAM CALCULATES PD AND PF
C
C       $A DTF2 2
C       PROGRAM CALLS SUBROUTINES PDPFM, IN, AND SKIP
C
        DOUBLE PRECISION PD,PF
        READ(5,5) ISKIP
5       FORMAT(I3)
        READ(5,10) N
10      FORMAT(I5)
        READ(5,15) PDO
15      FORMAT(F7.6)
1       READ(5,20) THRESH
        IF(ISKIP.NE.0) CALL SKIP(2,ISKIP)
20      FORMAT(F5.2)
        CALL PDPFM(THRESH,N,PDO,PD,PF)
        WRITE(6,30) PF,PD
30      FORMAT(2X,3HPF=,D13.6,5X,3HPD=,D13.6)
        REWIND 2
        GO TO 1
        STOP
        END
```

```
C       SUBROUTINE CALCULATES PF AND PD
C       FROM N VALUES OF L UNDER H0
C
C       $A DTF2 2
C       CALLS SUBROUTINE IN
C
        SUBROUTINE PDPFM(THRESH,N,PD0,PD,PF)
        DOUBLE PRECISION PD,PF,PROB
        PF=0.
        PD=PD0
        IXXX=251
        PROB=1./FLOAT(N)
        DO 10 I=1,N
        CALL IN(Y,IXXX)
        IF(Y.LT.THRESH) GO TO 10
        PF= PF + PROB
        PD= PD + Y*PROB
10      CONTINUE
        RETURN
        END
```

LIST OF REFERENCES

Adams, S. L. 1973. The Theory of Signal Detectability: Extension to Optimum Arrays. Technical Report No. 9, Adaptive Signal Detection Laboratory, Duke University, Durham, N.C.

Adams, S. L. and Nolte, L. W. 1975. Bayes Optimum Array Detection of Targets of Known Location. Accepted for publication in J. Acoust. Soc. Am.

Adams, S. L. and Nolte, L. W. 1975. Optimum Array Detection in Fluctuating Ambient Noise Fields. Accepted for publication in J. Acoust. Soc. Am.

Bartlett, M. S. 1951. An Inverse Matrix Adjustment Arising in Discriminant Analysis. Annals of Math. Statistics, Vol. 22, pp. 107-111.

Bienvenu, G. and Vernet, J. L. 1972. Enhancement of Antenna Performance by Adaptive Processing. Proceedings of the NATO Advanced Study Institute on Signal Processing with Emphasis on Underwater Acoustics. 21 August - 1 September, Loughborough, U.K.; also appeared in: Griffiths, J. W. R., Stocklin, P. L., and van Schooneveld, C. (ed.). 1973. Signal Processing (book). Academic Press, New York, N.Y., pp. 605-617.

Birdsall, T. G. 1968. Adaptive Detection Receivers and Reproducing Densities. Technical Report No. 194, Cooley Electronics Laboratory, University of Michigan, Ann Arbor, Michigan.

Birdsall, T. G. 1973. The Theory of Signal Detectability: ROC Curves and their Character. Technical Report No. 177, Cooley Electronics Laboratory, University of Michigan, Ann Arbor, Michigan.

Birdsall, T. G. and Gobien, J. O. 1973. Sufficient Statistics and Reproducing Densities in Simultaneous Sequential Detection and Estimation. IEEE Transactions on Information Theory, Vol. IT-19, No. 6, pp. 760-768.

Blachman, N. M. 1957. On Fourier Series for Gaussian Noise. Information and Control 1, pp. 56-63.

Bryn, F. 1962. Optimum Signal Processing of Three-Dimensional Arrays Operating on Gaussian Signals and Noise. J. Acoust. Soc. Am., Vol. 34, No. 3, pp. 289-297.

Chang, J. H. and Tuteur, F. B. 1971. A New Class of Adaptive Array Processors. J. Acoust. Soc. Am., Vol. 49, No. 3, pp. 639-649.

Cox, H. 1968. Interrelated Problems in Estimation and Detection, Parts I and II. Presented at NATO Advanced Study Institute on Signal Processing with Emphasis on Underwater Acoustics. 12-23 August, Enschede - The Netherlands.

Cox, H. 1969. Optimum Arrays and the Schwartz Inequality. J. Acoust. Soc. Am., Vol. 45, No. 1, pp. 228-232.

Gallop, M. A. 1971. Adaptive Optimum Array Detectors. Technical Report No. 7, Adaptive Signal Detection Laboratory, Duke University, Durham, N.C.

Gallop, M. A. and Nolte, L. W. 1974. Bayesian Detection of Targets of Unknown Location. IEEE Trans. on Aerospace and Electronic Systems, Vol. AES-10, No. 4, pp. 429-435.

Giraudon, C. 1972. Results on Active Sonar Optimum Array Processing. Proceedings of the NATO Advanced Study Institute on Signal Processing with Emphasis on Underwater Acoustics. 21 August - 1 September, Loughborough, U.K.; also appeared in: Griffiths, J. W. R., Stocklin, P. L., and van Schooneveld, C. (ed.). 1973. Signal Processing (book). Academic Press, New York, N.Y., pp. 495-505.

Goodman, N. R. 1963. Statistical Analysis Based on a Certain Multivariate Complex Gaussian Distribution (An Introduction). Annals of Math. Statistics, Vol. 34, pp. 152-177.

Hatsell, C. P. and Nolte, L. W. 1974. A New Generalized Likelihood Ratio Formula. IEEE Trans. on Systems, Man, and Cybernetics, Vol. SMC-4, No. 4, pp. 389-392.

Hodgkiss, W. S. and Nolte, L. W. 1975. A Note on the Calculation of Performance for Likelihood Ratio Processors via Computer Simulation. Accepted for publication in the IEEE Trans. on Information Theory.

Horton, C. W. 1969. Signal Processing of Underwater Acoustic Waves (book). U.S. Government Printing Office, Washington, D.C.

Jansma, D. 1969. The Theory of Signal Detectability: Bayesian Philosophy, Classical Statistics, and the Composite Hypothesis. Technical Report No. 200, Cooley Electronics Laboratory, University of Michigan, Ann Arbor, Michigan.

Marshall, R. A. 1973. Correlation of Fourier Coefficients in the Covariance Matrix. Senior Honors Paper, Dept. of Electrical Engineering, Duke University, Durham, N.C.

Mermoz, H. 1964. Filtrage Adapté et Utilization Optimale d'une Antenne. Proceedings of the NATO Advanced Study Institute on Signal Processing with Emphasis on Underwater Acoustics. 14-26 September, Grenoble, France.

Middleton, D. and Van Meter, D. 1954. Modern Statistical Approaches to Reception in Communication Theory. IRE Transactions, PGIT-4, pp. 119-145.

Middleton, D. and Groginsky, H. L. 1964. Detection of Random Acoustic Signals and Receivers with Distributed Elements: Optimum Receiver Structures for Normal Signal and Noise Fields. J. Acoust. Soc. Am., Vol. 36, No. 5, pp. 727-737.

Nolte, L. W. 1965. Adaptive Realizations of Optimum Detectors for Synchronous and Sporadic Recurrent Signals in Noise. Technical Report No. 163, Cooley Electronics Laboratory, University of Michigan, Ann Arbor, Michigan.

Nolte, L. W. 1966. An Adaptive Realization of the Optimum Receiver for a Synchronous Recurrent Waveform. IEEE Trans. on Information Theory, Vol. IT-12, No. 1, pp. 78-80.

Papoulis, A. N. 1965. Probability, Random Variables, and Stochastic Processes (book). McGraw-Hill Book Co., New York, N.Y.

Peterson, W. W., Birdsall, T. G., and Fox, W. C. 1954. The Theory of Signal Detectability. IRE Transactions, PGIT-4, pp. 161-211.

Roberts, R. A. 1965. On the Detection of a Signal Known Except for Phase. IEEE Trans. on Information Theory, Vol. IT-11, No. 1, pp. 76-82.

Schultheiss, P. M. 1968. Passive Sonar Detection in the Presence of Interference. J. Acoust. Soc. Am., Vol. 43, No. 3, pp. 418-425.

Stocklin, P. L. 1963. Space-Time Sampling and Likelihood Ratio Processing in Acoustic Pressure Fields. Jour. Brit. I.R.E., pp. 79-90.

Van Trees, H. L. 1968. Detection, Estimation, and Modulation Theory: Part I (book). John Wiley and Sons, Inc., New York, N.Y.

Young, G. O. and Howard, J. E. 1970. Antenna Processing for Surface Target Detection. IEEE Trans. on Antennas and Propagation, Vol. AP-18, No. 3, pp. 335-342.